UNDERSTANDING THE MOLECULAR MECHANISMS OF HAIR FOLLICLE STEM CELL QUIESCENCE AND GENOME PLASTICITY

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by

Jayhun Lee

August 2013

UNDERSTANDING THE MOLECULAR MECHANISMS OF HAIR FOLLICLE STEM CELL QUIESCENCE AND GENOME PLASTICITY

Jayhun Lee, Ph.D.
Cornell University 2013

Adult stem cells (SCs) utilize their abilities of self-renewal and differentiation to maintain proper homeostasis of their residing tissue. Malfunctioning of tissue SCs can result in multiple developmental disorders and cancers. To maintain balance between proliferation and quiescence, and faithfully execute fate determination, tissue SCs must be tightly regulated. Thus, understanding the molecular mechanisms of tissue SCs proliferation/quiescence and how their fate is determined are of critical importance in stem cell biology and medicine.

Hair follicle stem cells reside in the bulge and remain relatively quiescent throughout the cycling of the hair follicle. During quiescence, hair follicle stem cells choose between the two fates prior to activation – migrate out from their niche and become differentiated progenitors or remain in the niche and self-renew. To understand how 1) hair follicle stem cells quiescence is maintained and 2) their quiescence relates to fate decision, we explored different aspects of regulation.

In our effort to address the first question, we revealed how a key transcription factor of hair follicle stem cell activation and proliferation, Runx1, plays a role in regulating multiple Cyclin-Dependent Kinase inhibitors (CDKis). We found that multiple CDKis including p15, p21, p27, and p57 are highly expressed during quiescence. In the absence of p21, hair follicle stem cells fail to enter quiescence in a

timely manner. Moreover, we discovered that Runx1 and p21 interaction plays a context-dependent role under different proliferative environments. Finally, we discovered a novel function of CDKis in regulating transcription of other CDKis, independent of kinase-inhibitory function. Together, we unveil the robust mechanism of hair follicle stem cell quiescence by transcriptional regulation of cell cycle regulators.

To address the second question, we took a very different approach. Given the importance of histone methylation in regulating embryonic SCs self-renewal and differentiation, we asked how such regulation might take part in maintaining hair follicle stem cells quiescence and fate determination. We show that a distinct low level of different histone methylation marks characterizes quiescent hair follicle stem cells, and this level seems to be regulated by growth factor signaling including Bmp. These marks are associated with both transcriptional activation (H3K4me3) and repression (H3K9me3 and H3K27me3) of chromatin at specific gene promoters. Strikingly, globally low status of H3K4me3, H3K9me3, and H3K27me3 during quiescence is associated with higher genome plasticity, which likely allows hair follicle stem cells to make a fate decision prior to activation. Collectively, our findings suggest that the mechanism of maintaining hair follicle stem cells quiescence is robust and adds more biological significance to quiescence than previously realized by associating this state with plasticity for fate determination.

BIOGRAPHICAL SKETCH

Jayhun Lee was born in Phoenix, Arizona as a first child between his mother Aeju Kim and father Chang-Hee Lee. Jayhun moved to Chuncheon, a small city in South Korea when he was five and lived up to middle school years. In 2000, Jayhun entered Kangwon Science High School. He became very interested in biology and chemistry during his study. He was a leader of a student discussion group in chemistry, where his motivation in science grew significantly.

After two years of high school, Jayhun entered Korea Advanced Institute of Science and Technology (KAIST) in 2002. Majoring in Biological Sciences, he was introduced to and trained in molecular biology in the laboratories of Drs. Yeonsoo Seo and Gou Young Koh, where he not only learned how basic biological concepts are used to design and perform various experiments, but also learned about different aspects of molecular mechanisms of DNA replication and vascularization, respectively. These experiences motivated Jayhun to pursue molecular biology in depth in United States.

In August 2006, Jayhun entered Biochemistry, Molecular and Cell Biology (BMCB) program. He became very excited to face the biology of tissue stem cells in Dr. Tudorita Tumbar's laboratory during his rotation and joined the lab in 2007. Since then, he has pursued different aspects of regulation in hair follicle stem cell quiescence and its significance during tissue homeostasis. After graduation, Jayhun plans to stay in stem cell field and further explore the nature of stem cell biology and its application to medicine.

Dedicated to the One who saves

ACKNOWLEDGMENTS

No human being can live alone. I cannot be at where I am now without the support from others during my graduate study. First of all, I am truly and deeply grateful for my Ph.D. advisor Dr. Tumbar. She accepted me as her student, taught me how to design and perform experiments, and trained me with critical thinking skills in data analysis and discussions. Most importantly, her mentorship and passion toward biological research greatly inspired my attitude toward science.

I am also greatly thankful for all the guidance and useful advice I got from my committee members Drs. John Lis and Paul Soloway. Whenever I was troubled with experiments, they were supportive and provided with good suggestions that opened up the new paths. I also thank Dr. David Shalloway for advising and supporting of my PNAS paper.

I truly thank current and former colleagues of the Tumbar laboratory, Song Eun Lee, Karen Osorio, Ying Zhang, Cornelia Scheitz, Alex Wang, Aiko Sada, David McDermitt, Prachi Jain, Meng Zhang, Fadi Jacob, Tae Seung Lee, Shu Yang Lu, Chalrene Hoi and many others for all the support. I also thank the undergraduate students who worked under my direct supervision, Karin Lilja, Hye Yoon Song, and Keegan Colletier for assisting me with numerous experiments. They not only made the daily laboratory life more spirited, but also significantly contributed to both projects.

I also express great appreciation to my scientific colleagues outside the Tumbar laboratory. I thank Dr. Brian White for his significant contribution to my first

paper in mathematical and computational analysis of our data. I also thank Drs. Nasun Hah, Mike Guertin, Raga Krishnakumar for troubleshooting ChIP and ChIP-Seq experiments, Lavanya Gowri Sayam and James Lee Smith for the help with uncountable FACS experiments, Sylvie Allen for taking good care of my mice, and Drs. Lishuang Shen and Qi Sun for the help with ChIP-Seq data analysis. Their assistance was crucial in adding more significance to both projects.

I am deeply grateful for the care and relationship outside the laboratory. I can never thank enough to my church colleagues Pastor Seokjoon Chang and his family, Ki Suh Lee and his family, and all members of the praise team who supported me with prayer and love that allowed me to not give up and stand still by faith. I also thank Joon Woo Sohn and Sung-Jin Lee for their friendship and prayer support. Last but not least, thirty years of my life will not have existed without the support of my family. I cannot thank enough to my father, mother and my brother for their unconditional love and care. Thank you and I love you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
STEM CELL DYNAMICS DURING TISSUE HOMEOSTASIS

1.1 General notions of stem cells

Stem cells are functionally defined as cells capable of self-renewal and differentiation during an extended period of time (Smith 2001; Weissman, Anderson et al. 2001). Multiple stem cell types exist in an organism for proper homeostasis and they vary greatly in their innate and residential properties. The ability of stem cells to give rise to multiple lineages is termed 'potency' (Lanza 2004). Depending on the potency, stem cells can be hierarchically characterized as pluripotent, multipotent, oligopotent, and unipotent. Embryonic stem cells (ESCs) are derived from the inner cell mass of the blastocyst and are pluripotent, being able to give rise to all cell types of the organism – endodermal, ectodermal, and mesodermal lineages (Evans and Kaufman 1981; Martin 1981; Thomson, Itskovitz-Eldor et al. 1998). ESCs self-renew and are maintained *in vitro*. Many studies have identified critical regulators of ESC self-renewal such as Oct4, Sox2, Nanog and Myc. These factors have been recently used to reprogram already differentiated cells into pluripotent cells, which are named 'induced pluripotent stem cells (iPSCs)' (Takahashi and Yamanaka 2006). Very similar to ESCs, iPSCs carry pluripotency and recent studies in this field are focused

on efficient reprogramming as well as safer use of these cell types toward disease treatment (Cherry and Daley 2013).

Another type of stem cells is the tissue stem cells. Tissue stem cells are more lineage-restricted (oligopotent or multipotent) and they maintain the native tissue they reside in during normal adult homeostasis as well as help repair tissue injury (Morrison and Spradling 2008). Tissue stem cells are kept in a specialized microenvironment, known as the niche. There the tissue stem cells are provided with support and nurtured by the surroundings such as growth factors, vascularization, and other signaling molecules. Some tissue stem cells such as hematopoietic and hair follicle stem cells remain relatively quiescent within their niche, likely to keep away from stem cell exhaustion as well as accumulation of replication-associated genotoxic stress such as telomere shortening and DNA mutations (Fuchs 2009). On the other hand, other tissue stem cells such as epidermal and intestinal stem cells undergo rapid turnover during normal homeostasis suggesting different regulatory mechanisms exist in aiding tissue stem cell self-renewal and differentiation (Barker, van Es et al. 2007; Clayton, Doupe et al. 2007).

Stem cells have great potential for clinical applications due to their innate capability of being able to live long and give rise to differentiated cell types. Moreover, failure of proper functioning of stem cells very often results in multiple diseases such as cancer or tissue degeneration (Daley 2012). Thus, it is critical to understand the molecular basis and cellular behavior of stem cells for their further use in stem cell-based therapies. This chapter reviews various aspects of dynamic behavior and regulation of stem cells.

1.2 Hair follicle stem cells and adult hair cycle

Tissue stem cells possess the ability to maintain their native residing tissues intact during homeostasis (Fuchs 2009; Sada and Tumbar 2013). In order to do so, tissue stem cells balance between proliferation and quiescence as well as decisions on their fates. Different tissue stem cells utilizes different mechanisms such as symmetric or asymmetric cell division, as well as cell-extrinsic factors to maintain their stem cell pool while giving rise to cells that compose their native tissue (Sada and Tumbar 2013). Hair follicles undergo cycles of growth and degeneration during the life time of a mouse (Blanpain and Fuchs 2006; Lee and Tumbar 2012). Each cycle lasts about three weeks and consists of anagen (growth), catagen (regression) and telogen (quiescence) (Cotsarelis 2006). First, hair cycle starts around postnatal day (PD) 21 when hair follicles are at telogen. Prior to activation, some quiescent hair follicle stem cells migrate downward from the bulge (niche) to the hair germ without self-renewing, where cells become differentiated progenitors (Zhang, Cheong et al. 2009). The hair germ is a structure below the bulge that at least transiently contains stem cells. As cells of the hair germ start dividing rapidly, hair follicle is considered to be in anagen. During anagen, remaining cells of the bulge self-renew by symmetric divisions to maintain their pool. In addition, progenitors of the Matrix generate terminally differentiated cells that form different lineages of the Inner Root Sheath (IRS), as well as the hair shaft (Tumbar 2012). During catagen (~PD36), bulge cells re-enter quiescence, and the lower hair follicle structure undergoes apoptosis and regression.

The mouse hair follicle serves as an excellent model system for studying the dynamic behavior of tissue stem cells. First, the cyclic behavior is relatively

synchronous across the mouse skin, allowing researchers to isolate large quantities of stem cells at desired stages in their adult homeostatic cycle. Second, hair follicle stem cells reside in a distinct location (bulge) and their proliferative behavior directly correlates with the growth of hair follicles, which can be recognized by hair coat color, skin thickness, and histological approaches (Muller-Rover, Handjiski et al. 2001). Thirdly, molecular characterization has revealed that hair follicle stem cells possess distinct molecular makeup, providing researchers a basis for understanding different regulatory axes of hair follicle stem cell dynamics (Tumbar, Guasch et al. 2004). Lastly, hair follicle stem cells act as a pool in which a few of them migrate out from the niche while they are still quiescent, and are destined to become progenitors that form the mature hair follicle (Zhang, Cheong et al. 2009). Only later the stem cells remaining in the niche undergo self-renewal to maintain their stem cell pool. This provides a temporal separation of two important mechanisms – proliferation (anagen) and fate choice (catagen-telogen) – allowing us to dissect both aspects using hair follicle stem cells. Importantly, many of the findings made from using hair follicle as a model can be applied to other tissue stem cell systems as well.

Classically, before bulge cells were shown to contain hair follicle stem cells, bulge cells were denoted as label-retaining cells (LRCs) due to their infrequently dividing property. The initial labeling with 3[H]-thymidine or 5-Bromo-deoxyuridine (BrdU), nucleotide analogs that get incorporated in the DNA during S phase, was retained over time (Cotsarelis, Sun et al. 1990; Taylor, Lehrer et al. 2000; Oshima, Rochat et al. 2001). From the observation that these bulge LRCs contribute to other regions of hair follicles, subsequent studies revealed that the bulge LRCs carry

multipotent stem cells. Single cells isolated from the bulge were able to be cultured *in vitro* for a prolonged time, and were still able to contribute to all lineages of hair follicle after two consecutive transplantations (Claudinot, Nicolas et al. 2005) as well as long-term lineage tracings (Zhang, Cheong et al. 2009).

Recently, more precise mapping of hair follicle stem cells proliferation dynamics was done using the tetracycline-inducible (tet-off) system (Tumbar, Guasch et al. 2004; Zhang, Cheong et al. 2009; Zhang, White et al. 2010). When doxycycline is introduced to mouse diet, H2B-GFP mRNA expression is shut off and upon division, the H2B-GFP protein is equally diluted in half in between daughter cells (Waghmare, Bansal et al. 2008). These studies revealed that the bulge cells on average divide three times during one hair cycle. Chapter 2 contains a concise review of dynamic regulatory mechanisms such as signaling molecules, growth factors, key transcription factors, and hair follicle surrounding environments that play critical roles in hair follicle stem cells self-renewal and differentiation (Lee and Tumbar 2012).

1.3 Cell cycle regulation of stem cells

Progression through the cell-cycle depends on levels of Cyclin/CDK (Cyclin-Dependent Kinase). For instance, level of Cyclin D/CDK4 and Cyclin D/CDK6 complexes peak at G1 to S transition which allow further progression into DNA replication. The activities of CDKs are blocked by CDK-inhibitors (CDKis). CDKis consists of two families – a Cip/Kip family (p21, p27, and p57) and an Ink4 family

(p15, p16, p18, and p19) (Besson, Dowdy et al. 2008). The Ink4 family of the CDKis inhibit CDK4 and CDK6 during G1 to S transition, which results in cell-cycle dependent behavior. On the other hand, Cip/Kip family of CDKis show broader spectrum of inhibition throughout the cell cycle (Tesio and Trumpp 2011).

Although it has been recently proposed that some tissue stem cells may actively cycle throughout life (Li and Clevers 2010), it is generally believed that tight regulation of quiescence is fundamental in maintaining tissue stem cells and in preventing cancer (Orford and Scadden 2008; Fuchs 2009). Cyclin-Dependent Kinase inhibitors (CDKis) play specific roles in regulating cell division in different tissue cell types. The Cip/Kip family (p21, p27, and p57) and the Ink4 family (p15, p16, p18, and p19) of CDKis have been studied in the hematopoietic stem cell system, where they may maintain quiescence and prevent stem cell exhaustion (Matsumoto, Takeishi et al. 2011; Tesio and Trumpp 2011; Zou, Yoshihara et al. 2011). However, cell cycle regulation in other tissue stem cells, such as those of skin and hair follicle, is understudied.

Signaling pathways converge onto transcription factors that may ultimately interact with CDKis and control their mRNA expression. However, molecular interactions controlling quiescence in tissue stem cells, and how they differ in normal homeostasis when compared to a context in which the niche influence on stem cells has been perturbed is poorly understood.

An interesting case of stem cell cell-cycle control is that of p21 (*Cdkn1a*), which seems to inhibit stem cell proliferation in cell culture or transplantation in three tissue types - skin (Topley, Okuyama et al. 1999), blood (Cheng, Rodrigues et al. 2000;

Viale, De Franco et al. 2009) and brain (Kippin, Martens et al. 2005). However, controversial findings show that p21 may have limited roles in normal blood stem cell homeostasis (van Os, Kamminga et al. 2007; Foudi, Hochedlinger et al. 2009; Matsumoto, Takeishi et al. 2011; Zou, Yoshihara et al. 2011) and instead only act in injury conditions.

In skin and hair follicle, the role of p21 in normal homeostasis has not been addressed extensively yet. In addition, we previously found that p21 and the transcription factor Runx1, a Runt family member of cancer genes (Speck and Gilliland 2002) and a regulator of hair follicle stem cell proliferation (Osorio, Lee et al. 2008), interact genetically in cultured hair follicle stem cells (Hoi, Lee et al. 2010) but the significance of this interaction in the normal tissue also remains unclear.

p21 knockout mice display no observable defects. However, absence of p21 led to an increased short-term proliferation potential of *in vitro* epithelial-derived skin keratinocytes as well as increased susceptibility to skin carcinoma formation by chemical induction *in vivo* (Topley, Okuyama et al. 1999). p21, initially identified as a transcriptional target of the tumor suppressor p53, plays a role in various cellular processes including stem cell aging and maintenance of stem cell quiescence (Weinberg and Denning 2002; Ju, Choudhury et al. 2007; Boyer and Cheng 2008). This is accomplished by its ability to block S-phase progression by binding to proliferating cell nuclear antigen (PCNA), a subunit required for proper DNA replication. In addition to its role during normal homeostasis, p21 level is elevated upon DNA damage in response to p53 up-regulation, resulting in cell cycle arrest or apoptosis (Weinberg and Denning 2002). The level and localization of p21 is

important in making a choice between cell death and cell cycle arrest. An apoptotic executer, caspase-3 uses p21 as its substrate and cleaves the C-terminal residues including the nuclear localization signal and PCNA-interacting domain, compromising its activity and localization (Suzuki, Tsutomi et al. 1998; Zhou, Liao et al. 2001). Loss of p21 can result in apoptotic cells at areas of neovascularization (Bruhl, Heeschen et al. 2004).

Recently, in addition to their roles in cell cycle regulation, some CDKis (p21 and p27) have also been shown to play cell cycle-independent roles and can function as transcriptional regulators of other genes in conjunction with transcription factors such as E2F (Devgan, Mammucari et al. 2005; Pippa, Espinosa et al. 2011; Ferrandiz, Caraballo et al. 2012). In Chapter 3, we show how Runx1 transcriptionally regulates multiple CDKis mRNA expression in hair follicle stem cells. Moreover, we address the role of one CDKi (p21) in regulating timely exit from the cell cycle. In addition, the significance and context-dependent role of Runx1-p21 interaction is analyzed. Finally, we show a novel mechanism of transcriptional regulation of one CDKi (p15) by another CDKi (p21) (Lee, Hoi et al. 2013).

1.4 Histone modifications and their regulation of stem cell function

Chromatin is a complex genetic material of DNA packaged with histones and non-histone proteins that form the chromosome (Luger, Mader et al. 1997; Luger, Dechassa et al. 2012). The basic organizational unit of chromatin is the nucleosome, which consists of 4 core histones (H2A, H2B, H3, and H4) wrapped around ~147bp of

DNA, known as “10nm-fiber” or “beads-on-a-string” as visualed by electron-microscopy (Luger, Mader et al. 1997). Chromatin further compacts into higher order domains, forming transcriptionally active “euchromatin” or silent “heterochromatin” (Hansen 2002; Horn and Peterson 2002; Giadrossi, Dvorkina et al. 2007).

Different transcriptional status is driven not only by transcription factors but also by multiple factors such as DNA methylation, histone variants, chromatin remodeling factors, histone N-terminal modifications etc (Jenuwein and Allis 2001; Giadrossi, Dvorkina et al. 2007). Chromatin modifications are relatively well established among these factors (Jenuwein and Allis 2001; Kouzarides 2007). For instance, H3K4me3 (histone 3 lysine 4 tri-methylation) is associated with active genes, while H3K27me3 and H3K9me3 are associated with silenced genes (Bernstein, Mikkelsen et al. 2006). These modifications are placed and maintained by specific histone modifying enzymes. Suv39h1/h2 (Kmt1a/1b) containing SET (enzymatic functional domain of methyl-transferase activity) domain selectively methylates K9 position (Rea, Eisenhaber et al. 2000; Lehnertz, Ueda et al. 2003). On the other hand, histone demethylases such as Jmjd2a (Kdm4a), Jmjd1a (Kdm3a), and Jmjd2c (Kdm4c) can remove methyl- residues on same position (Marmorstein and Trievel 2009). H3K27me3 is placed on and maintained by Ezh2 (Kmt6), an enzymatic component of Polycomb Repressive Complexes and can be removed by demethylases such as Jmjd3 (Kdm6b), and Utx (Kdm6a) (Shen, Kim et al. 2009; Margueron and Reinberg 2011; Orkin and Hochedlinger 2011). Set1a/1b as well as MLL1-4 are homologs of Drosophila Trithorax group complexes and function in placing and maintaining H3K4me3 mark on active genes, while another set of demethylases act

specifically to remove this mark (Roguev, Schaff et al. 2001; Miller, Krogan et al. 2001; Milne, Briggs et al. 2002; Iwase, Lan et al. 2007) . Recent studies revealed that some developmental genes are poised in their expressions via presence of both active (H3K4me3) and repressive (H3K27me3) marks termed 'bivalency', which is resolved upon lineage restriction (Bernstein, Mikkelsen et al. 2006; Sparmann and van Lohuizen 2006; Mikkelsen, Ku et al. 2007).

Together with genetic studies of histone modifying enzymes and next-generation sequencing technology, histone modification maps and their regulatory networks in different cell types including stem cells, lineage-restricted cells, and cancer cells have revealed the great importance of epigenetic regulations in stem cell biology. The chromatin of embryonic stem cells is more open and hyperdynamic, and differentiation reduces the chromatin dynamics (Meshorer, Yellajoshula et al. 2006). In addition, embryonic stem cells global chromatin composition is distinct from that of differentiated cell types (Hawkins, Hon et al. 2010). Moreover, core pluripotency factors such as Oct4, Sox2, Klf4 and Myc (Takahashi and Yamanaka 2006) that are critical for reprogramming as well as other factors such as Nanog, interact with multiple chromatin factors including remodelers and histone modifying enzymes that lead to global changes to the gene expression, which determines cell fate (Orkin and Hochedlinger 2011). Interestingly, recent studies revealed that histone modifying enzymes act as important factors in mediating efficient reprogramming (Onder, Kara et al. 2012) by transcription factors and that H3K9 methylation hinders the reprogramming process (Chen, Liu et al. 2013).

Specific manipulations of Polycomb group proteins as well as different histone modifying enzymes have revealed that tissue stem cells also employ multiple mechanisms of chromatin regulation by histone methylation in not only maintaining the stem cell pool, but also in proper differentiation, tissue homeostasis and cancer (Park, Qian et al. 2003; Iwama, Oguro et al. 2004; Molofsky, Slutsky et al. 2006; Fasano, Dimos et al. 2007; Sangiorgi and Capecchi 2008; Sen, Webster et al. 2008; Ezhkova, Pasolli et al. 2009; Ezhkova, Lien et al. 2011; Margueron and Reinberg 2011; Chen, Skutt-Kakaria et al. 2012; Ye, Fan et al. 2012; Pal, Bouras et al. 2013). However, the consequences of perturbing these modifications are shockingly mild, most likely due to compensation or potential redundancy among multiple histone marks, which we are just beginning to understand. Recent studies suggest that Ezh2-mediated changes in H3K27me3 govern the rates and timing of cell differentiation in the mammary epithelium (Pal, Bouras et al. 2013). In addition, Ezh2 fine tunes the timing of developmental switches and the balance between self-renewal and differentiation of multiple stem/progenitors including, skeletal (Juan, Derfoul et al. 2011), hematopoietic (Su, Basavaraj et al. 2003; Mochizuki-Kashio, Mishima et al. 2011), neural (Pereira, Sansom et al. 2010). Moreover, another subunit of Polycomb Repressive Complex, Bmi1, also plays critical role in similar processes (Park, Qian et al. 2003; Iwama, Oguro et al. 2004; Fasano, Dimos et al. 2007; Sangiorgi and Capecchi 2008).

Some histone demethylases (Kdm4b and Kdm6b) are regulated by BMP signaling, which is critical for osteogenic differentiation of mesenchymal stem cells (Ye, Fan et al. 2012). Importantly, mis-regulation of factors implicated in histone

methylation is commonly found in cancers, suggesting that balancing the levels of these class of epigenetic regulators is essential not only for normal homeostasis but also for preventing cancer (Sauvageau and Sauvageau 2010).

Proliferation associated changes in chromatin organization were also observed not only in tissue stem cells but also in more lineage-committed cells. For instance, histone hypomethylation was a distinct characteristic of quiescent B lymphocytes, which was restored upon proliferation onset (Baxter, Sauer et al. 2004). In addition, these quiescent B lymphocytes were more efficiently reprogrammed upon somatic nuclear transfer suggesting a higher plasticity associated with quiescence (Baxter, Sauer et al. 2004).

In skin, Ezh2-deficient basal layer cells of the epidermis undergo precocious epidermal differentiation during skin development (Ezhkova, Pasolli et al. 2009; Ezhkova, Lien et al. 2011). Ezh2 is enriched in basal layer cells and its loss results in decreased proliferation of epidermal keratinocytes *in vivo* as well as *in vitro*. Repression through H3K27me3 on differentiation-specific genes is relieved upon Ezh2 loss and results in precocious expression of late-differentiation genes and acquisition of skin barrier function. Postnatal loss of Ezh2 is compensated by Ezh1, and loss of both appears to impair proliferation of hair follicle progenitor matrix cells and lack of differentiated lineages during hair morphogenesis (Ezhkova, Lien et al. 2011). Moreover, recently it has been shown that complete lack of H3K27me3 via double knockout of both Ezh1 and Ezh2 in epidermis progenitors results in loss of a specific lineage specialized in mechano-sensing of fine touch (Ezhkova, Lien et al. 2011; Bardot, Valdes et al. 2013). This was in fact the first and only example to date

in which H3K27me3 proved indispensable for specification of a differentiated lineage, but also skin is the only tissue where both enzymes have been knocked out simultaneously. Although work in embryonic SCs predicted a more widespread role of the H3K27me3 mark in tissue development, since Ezh1 and Ezh2 act redundantly, it remains to be seen if knocking them out together in other tissues besides skin will broaden our understanding of their function in lineage specifications.

Absence of Jmjd3, an H3K27me3-demethylase, blocks differentiation of organotypic human epidermal cells while its over-expression induces differentiation (Sen, Webster et al. 2008). However, if and how adult hair follicle stem cells employ histone modifications as means of regulation to achieve proper hair follicle homeostasis require further exploration.

In Chapter 4, we show preliminary results of regulation of histone H3 tri-methylations at K9, K27, and K4 that takes place during hair follicle homeostasis. We show that hair follicle stem cells have a distinct landscape of these histone modifications in their genome that undergoes global changes upon self-renewal and differentiation processes. These changes seem to be driven by transcriptional regulation of multiple histone modifying enzymes triggered by one or more growth factor signaling. Finally, the global status of quiescent hair follicle stem cells seems to place the chromatin in a more hyperdynamic state as suggested by histone H2B-GFP exchange rates. Taken together, our data suggest a model in which the state of hair follicle stem cells in quiescence is associated with low levels of both repressive and activating histone marks and more hyper-dynamic chromatin which may reflect low epigenomic identity and higher genome plasticity. This state may be important for the

fate choice between self-renewal and differentiation that takes place in quiescence for hair follicle and potentially for other tissue stem cells. It will remain to be shown whether this epigenetic state is crucial for fate decisions, and whether other epigenetic marks including other histone modifications may follow the same patterns as the ones we describe here.

REFERENCES

Bardot, E. S., V. J. Valdes, et al. (2013). "Polycomb subunits Ezh1 and Ezh2 regulate the Merkel cell differentiation program in skin stem cells." EMBO J.

Barker, N., J. H. van Es, et al. (2007). "Identification of stem cells in small intestine and colon by marker gene Lgr5." Nature 449(7165): 1003-1007.

Baxter, J., S. Sauer, et al. (2004). "Histone hypomethylation is an indicator of epigenetic plasticity in quiescent lymphocytes." EMBO J 23(22): 4462-4472.

Bernstein, B. E., T. S. Mikkelsen, et al. (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell 125(2): 315-326.

Besson, A., S. F. Dowdy, et al. (2008). "CDK inhibitors: cell cycle regulators and beyond." Dev Cell 14(2): 159-169.

Blanpain, C. and E. Fuchs (2006). "Epidermal stem cells of the skin." Annu Rev Cell Dev Biol 22: 339-373.

Boyer, M. J. and T. Cheng (2008). "The CDK inhibitors: potential targets for therapeutic stem cell manipulations?" Gene Ther 15(2): 117-125.

Bruhl, T., C. Heeschen, et al. (2004). "p21Cip1 levels differentially regulate turnover of mature endothelial cells, endothelial progenitor cells, and in vivo neovascularization." Circ Res 94(5): 686-692.

Chen, J., H. Liu, et al. (2013). "H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs." Nat Genet 45(1): 34-42.

Chen, X., K. Skutt-Kakaria, et al. (2012). "G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment." Genes Dev 26(22): 2499-2511.

Cheng, T., N. Rodrigues, et al. (2000). "Hematopoietic stem cell quiescence maintained by p21cip1/waf1." Science 287(5459): 1804-1808.

Cherry, A. B. and G. Q. Daley (2013). "Reprogrammed cells for disease modeling and regenerative medicine." Annu Rev Med 64: 277-290.

Claudinot, S., M. Nicolas, et al. (2005). "Long-term renewal of hair follicles from clonogenic multipotent stem cells." Proc Natl Acad Sci U S A 102(41): 14677-14682.

Clayton, E., D. P. Doupe, et al. (2007). "A single type of progenitor cell maintains normal epidermis." Nature 446(7132): 185-189.

Cotsarelis, G. (2006). "Epithelial stem cells: a folliculocentric view." J Invest Dermatol 126(7): 1459-1468.

Cotsarelis, G., T. T. Sun, et al. (1990). "Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis." Cell 61(7): 1329-1337.

Daley, G. Q. (2012). "The promise and perils of stem cell therapeutics." Cell Stem Cell 10(6): 740-749.

Devgan, V., C. Mammucari, et al. (2005). "p21WAF1/Cip1 is a negative transcriptional regulator of Wnt4 expression downstream of Notch1 activation." Genes Dev 19(12): 1485-1495.

Evans, M. J. and M. H. Kaufman (1981). "Establishment in culture of pluripotential cells from mouse embryos." Nature 292(5819): 154-156.

Ezhkova, E., W. H. Lien, et al. (2011). "EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair." Genes Dev 25(5): 485-498.

Ezhkova, E., H. A. Pasolli, et al. (2009). "Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells." Cell 136(6): 1122-1135.

Fasano, C. A., J. T. Dimos, et al. (2007). "shRNA knockdown of Bmi-1 reveals a critical role for p21-Rb pathway in NSC self-renewal during development." Cell Stem Cell 1(1): 87-99.

Ferrandiz, N., J. M. Caraballo, et al. (2012). "p21 as a Transcriptional Co-Repressor of S-Phase and Mitotic Control Genes." PLoS One 7(5): e37759.

Foudi, A., K. Hochedlinger, et al. (2009). "Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells." Nat Biotechnol 27(1): 84-90.

Fuchs, E. (2009). "The tortoise and the hair: slow-cycling cells in the stem cell race." Cell 137(5): 811-819.

Giadrossi, S., M. Dvorkina, et al. (2007). "Chromatin organization and differentiation in embryonic stem cell models." Curr Opin Genet Dev 17(2): 132-138.

Hansen, J. C. (2002). "Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions." Annu Rev Biophys Biomol Struct 31: 361-392.

Hawkins, R. D., G. C. Hon, et al. (2010). "Distinct epigenomic landscapes of pluripotent and lineage-committed human cells." Cell Stem Cell 6(5): 479-491.

Hoi, C. S., S. E. Lee, et al. (2010). "Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin." Mol Cell Biol 30(10): 2518-2536.

Horn, P. J. and C. L. Peterson (2002). "Molecular biology. Chromatin higher order folding--wrapping up transcription." Science 297(5588): 1824-1827.

Iwama, A., H. Oguro, et al. (2004). "Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1." Immunity 21(6): 843-851.

Iwase, S., F. Lan, et al. (2007). "The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases." Cell 128(6): 1077-1088.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science 293(5532): 1074-1080.

Ju, Z., A. R. Choudhury, et al. (2007). "A dual role of p21 in stem cell aging." Ann N Y Acad Sci 1100: 333-344.

Juan, A. H., A. Derfoul, et al. (2011). "Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells." Genes Dev 25(8): 789-794.

Kippin, T. E., D. J. Martens, et al. (2005). "p21 loss compromises the relative quiescence of forebrain stem cell proliferation leading to exhaustion of their proliferation capacity." Genes Dev 19(6): 756-767.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell 128(4): 693-705.

Lanza, R. P. (2004). Handbook of stem cells. Boston, MA, Elsevier Academic.

Lee, J., C. S. Hoi, et al. (2013). "Runx1 and p21 synergistically limit the extent of hair follicle stem cell quiescence in vivo." Proc Natl Acad Sci U S A 110(12): 4634-4639.

Lee, J. and T. Tumbar (2012). "Hairy tale of signaling in hair follicle development and cycling." Semin Cell Dev Biol 23(8): 906-916.

Lehnertz, B., Y. Ueda, et al. (2003). "Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin." Curr Biol 13(14): 1192-1200.

Li, L. and H. Clevers (2010). "Coexistence of quiescent and active adult stem cells in mammals." Science 327(5965): 542-545.

Luger, K., M. L. Dechassa, et al. (2012). "New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?" Nat Rev Mol Cell Biol 13(7): 436-447.

Luger, K., A. W. Mader, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature 389(6648): 251-260.

Margueron, R. and D. Reinberg (2011). "The Polycomb complex PRC2 and its mark in life." Nature 469(7330): 343-349.

Marmorstein, R. and R. C. Trievel (2009). "Histone modifying enzymes: structures, mechanisms, and specificities." Biochim Biophys Acta 1789(1): 58-68.

Martin, G. R. (1981). "Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells." Proc Natl Acad Sci U S A 78(12): 7634-7638.

Matsumoto, A., S. Takeishi, et al. (2011). "p57 is required for quiescence and maintenance of adult hematopoietic stem cells." Cell Stem Cell 9(3): 262-271.

Meshorer, E., D. Yellajoshula, et al. (2006). "Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells." Dev Cell 10(1): 105-116.

Mikkelsen, T. S., M. Ku, et al. (2007). "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells." Nature 448(7153): 553-560.

Miller, T., N. J. Krogan, et al. (2001). "COMPASS: a complex of proteins associated with a trithorax-related SET domain protein." Proc Natl Acad Sci U S A 98(23): 12902-12907.

Milne, T. A., S. D. Briggs, et al. (2002). "MLL targets SET domain methyltransferase activity to Hox gene promoters." Mol Cell 10(5): 1107-1117.

Mochizuki-Kashio, M., Y. Mishima, et al. (2011). "Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells." Blood 118(25): 6553-6561.

Molofsky, A. V., S. G. Slutsky, et al. (2006). "Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing." Nature 443(7110): 448-452.

Morrison, S. J. and A. C. Spradling (2008). "Stem cells and niches: mechanisms that promote stem cell maintenance throughout life." Cell 132(4): 598-611.

Muller-Rover, S., B. Handjiski, et al. (2001). "A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages." J Invest Dermatol 117(1): 3-15.

Onder, T. T., N. Kara, et al. (2012). "Chromatin-modifying enzymes as modulators of reprogramming." Nature 483(7391): 598-602.

Orford, K. W. and D. T. Scadden (2008). "Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation." Nat Rev Genet 9(2): 115-128.

Orkin, S. H. and K. Hochedlinger (2011). "Chromatin connections to pluripotency and cellular reprogramming." Cell 145(6): 835-850.

Oshima, H., A. Rochat, et al. (2001). "Morphogenesis and renewal of hair follicles from adult multipotent stem cells." Cell 104(2): 233-245.

Osorio, K. M., S. E. Lee, et al. (2008). "Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation." Development 135(6): 1059-1068.

Pal, B., T. Bouras, et al. (2013). "Global changes in the mammary epigenome are induced by hormonal cues and coordinated by ezh2." Cell Rep 3(2): 411-426.

Park, I. K., D. Qian, et al. (2003). "Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells." Nature 423(6937): 302-305.

Pereira, J. D., S. N. Sansom, et al. (2010). "Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex." Proc Natl Acad Sci U S A 107(36): 15957-15962.

Pippa, R., L. Espinosa, et al. (2011). "p27(Kip1) represses transcription by direct interaction with p130/E2F4 at the promoters of target genes." Oncogene.

Rea, S., F. Eisenhaber, et al. (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature 406(6796): 593-599.

Roguev, A., D. Schaft, et al. (2001). "The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4." EMBO J 20(24): 7137-7148.

Sada, A. and T. Tumbar (2013). "New insights into mechanisms of stem cell daughter fate determination in regenerative tissues." Int Rev Cell Mol Biol 300: 1-50.

Sangiorgi, E. and M. R. Capecchi (2008). "Bmi1 is expressed in vivo in intestinal stem cells." Nat Genet 40(7): 915-920.

Sauvageau, M. and G. Sauvageau (2010). "Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer." Cell Stem Cell 7(3): 299-313.

Sen, G. L., D. E. Webster, et al. (2008). "Control of differentiation in a self-renewing mammalian tissue by the histone demethylase JMJD3." Genes Dev 22(14): 1865-1870.

Shen, X., W. Kim, et al. (2009). "Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells." Cell 139(7): 1303-1314.

Smith, A. G. (2001). "Embryo-derived stem cells: of mice and men." Annu Rev Cell Dev Biol 17: 435-462.

Sparmann, A. and M. van Lohuizen (2006). "Polycomb silencers control cell fate, development and cancer." Nat Rev Cancer 6(11): 846-856.

Speck, N. A. and D. G. Gilliland (2002). "Core-binding factors in haematopoiesis and leukaemia." Nat Rev Cancer 2(7): 502-513.

Su, I. H., A. Basavaraj, et al. (2003). "Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement." Nat Immunol 4(2): 124-131.

Suzuki, A., Y. Tsutomi, et al. (1998). "Resistance to Fas-mediated apoptosis: activation of caspase 3 is regulated by cell cycle regulator p21WAF1 and IAP gene family ILP." Oncogene 17(8): 931-939.

Takahashi, K. and S. Yamanaka (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." Cell 126(4): 663-676.

Taylor, G., M. S. Lehrer, et al. (2000). "Involvement of follicular stem cells in forming not only the follicle but also the epidermis." Cell 102(4): 451-461.

Tesio, M. and A. Trumpp (2011). "Breaking the cell cycle of HSCs by p57 and friends." Cell Stem Cell 9(3): 187-192.

Thomson, J. A., J. Itskovitz-Eldor, et al. (1998). "Embryonic stem cell lines derived from human blastocysts." Science 282(5391): 1145-1147.

Topley, G. I., R. Okuyama, et al. (1999). "p21(WAF1/Cip1) functions as a suppressor of malignant skin tumor formation and a determinant of keratinocyte stem-cell potential." Proc Natl Acad Sci U S A 96(16): 9089-9094.

Tumbar, T. (2012). "Ontogeny and Homeostasis of Adult Epithelial Skin Stem Cells." Stem Cell Reviews and Reports 8(2): 561-576.

Tumbar, T., G. Guasch, et al. (2004). "Defining the epithelial stem cell niche in skin." Science 303(5656): 359-363.

van Os, R., L. M. Kamminga, et al. (2007). "A Limited role for p21Cip1/Waf1 in maintaining normal hematopoietic stem cell functioning." Stem Cells 25(4): 836-843.

Viale, A., F. De Franco, et al. (2009). "Cell-cycle restriction limits DNA damage and maintains self-renewal of leukaemia stem cells." Nature 457(7225): 51-56.

Waghmare, S. K., R. Bansal, et al. (2008). "Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells." EMBO J 27(9): 1309-1320.

Weinberg, W. C. and M. F. Denning (2002). "P21Waf1 control of epithelial cell cycle and cell fate." Crit Rev Oral Biol Med 13(6): 453-464.

Weissman, I. L., D. J. Anderson, et al. (2001). "Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations." Annu Rev Cell Dev Biol 17: 387-403.

Ye, L., Z. Fan, et al. (2012). "Histone demethylases KDM4B and KDM6B promotes osteogenic differentiation of human MSCs." Cell Stem Cell 11(1): 50-61.

Zhang, Y. V., J. Cheong, et al. (2009). "Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells." Cell Stem Cell 5(3): 267-278.

Zhang, Y. V., B. S. White, et al. (2010). "Stem cell dynamics in mouse hair follicles: a story from cell division counting and single cell lineage tracing." Cell Cycle 9(8): 1504-1510.

Zhou, B. P., Y. Liao, et al. (2001). "Cytoplasmic localization of p21Cip1/WAF1 by Akt-induced phosphorylation in HER-2/neu-overexpressing cells." Nat Cell Biol 3(3): 245-252.

Zou, P., H. Yoshihara, et al. (2011). "p57(Kip2) and p27(Kip1) cooperate to maintain hematopoietic stem cell quiescence through interactions with Hsc70." Cell Stem Cell 9(3): 247-261.

CHAPTER 2

HAIRY TALE OF SIGNALING IN HAIR FOLLICLE DEVELOPMENT AND CYCLING [1]

2.1 Summary

Hair follicle (HF) is an appendage from the vertebrate skin epithelium, and is critical for environmental sensing, animal appearance, and body heat maintenance. HFs arise from the embryonic ectoderm and regenerate cyclically during adult life. Distinct morphological and functional stages from development through homeostasis have been extensively studied for the past decades to dissect the critical molecular mechanisms. Accumulating work suggests that different signaling cascades, such as Wnt, Bmp, Shh, and Notch, together with specific combinations of transcription factors are at work at different stages. Here we provide a comprehensive review of mouse genetics studies, which include lineage tracing along with knockout and over-expression of core genes from key signaling pathways, to paint an updated view of the molecular regulatory network that govern each stage of hair follicle development and adult cycling.

2.2 Introduction to general notions of cell and tissue signaling

[1] This chapter has been published: Lee, J. and Tumbar, T (2012). "Hairy tale of signaling in hair follicle development and cycling." Semin Cell Dev Biol **23**(08): 906-916.

Proper tissue architecture and function depends upon each cell in the body faithfully fulfilling the purpose of its existence. The harmony among complex signaling networks within and between cells must be strictly maintained in order to prevent developmental defects and fatal diseases. Skin homeostasis is essential for survival, and protects the organism against various environmental assaults such as UV irradiation, microbes, harmful chemicals, etc. (Blanpain and Fuchs 2006). Mammalian skin epithelium carries hair follicle (HF) appendages important for overall skin homeostasis, sensory purposes, maintaining body temperature, and establishing proper appearance. Moreover, HF is an ideal mammalian model system to understand more general principles of intra- and inter-cell signaling pathway communication that may apply to other tissues. HF possesses distinct characteristics that allow in depth dissection of molecular signaling. First, synchronous, distinct and dramatic morphological changes allow researchers to easily identify well-characterized HF stages of development or homeostasis (Muller-Rover, Handjiski et al. 2001; Blanpain and Fuchs 2006). Secondly, hair follicle stem cells (HFSCs) of the bulge, which give rise to all HF lineages and maintain homeostasis, are well characterized (Cotsarelis, Sun et al. 1990; Hsu and Fuchs 2012). HFSCs dynamic behavior of self-renewal, differentiation, and quiescence, is essential for the HF proliferative and growth cycle (Cotsarelis 2006; Zhang, Cheong et al. 2009). Lastly, stem, progenitor, and terminally differentiated cells display specific HF location and molecular marker expression (Hsu and Fuchs 2012), which allows their identification, targeting and isolation for deeper exploration.

The signals that govern skin and HF cell proliferation and differentiation during development and homeostasis have been intensively explored in the past decades (Table 2.1) (Plikus 2012). Several transcription factors such as Runx1 (Osorio, Lee et al. 2008; Hoi, Lee et al. 2010), Tcf3/4 (Nguyen, Merrill et al. 2009), NFATc1 (Horsley, Aliprantis et al. 2008), Lhx2 (Rhee, Polak et al. 2006), Sox9 (Vidal, Chaboissier et al. 2005), Tbx1 (Chen, Heller et al. 2012), p63 (Mills, Zheng et al. 1999; Yang, Schweitzer et al. 1999), Dlx3 (Hwang, Mehrani et al. 2008), Lef1 (Vangenderen, Okamura et al. 1994), Msx2/Foxn1 (Cai, Lee et al. 2009), Gli (Grachtchouk, Mo et al. 2000; Mill, Mo et al. 2003), Gata3 (Kaufman, Zhou et al. 2003), and etc., play critical roles in HFSC activation, self-renewal, differentiation, and cycling by modulating several key signaling pathways, which include Wnt, Bmp, Shh, and Notch. We begin by underlying basic molecular principles for each pathway, and next we review how different combinations of these pathways regulate HF development and cycling.

2.2.1 Wnt signaling

Wnt molecules act as secreted, morphogenetic signals, which bind to Frizzled or LRPs receptors (Gumbiner 1997; Alonso and Fuchs 2003; Logan and Nusse 2004). When Wnt is absent, Axin/GSK-3β /APC kinase complex phosphorylates serine/threonine residues of cytoplasmic β-catenin, leading to ubiquitin-mediated degradation. Upon Wnt ligand-binding to receptors, Dishevelled binds to the intra-cellular domains of the receptors, sequestering and allowing LRP-mediated phosphorylation of Axin/GSK-3β/APC complex, which inhibits phosphorylation and

Table 2.1. Categorized genetic studies of key components of signaling pathway and transcription factors in hair follicles.

Signals	Mutation type	Hair follicle phenotype	Reference	
1. Wnt-related				
Wnt3	*K14-Wnt3*	Reduced hair follicle length, premature catagen onset	Millar (1999)	
Wnt7a	*K14-Wnt7a*	Increased hair follicle neogenesis upon wounding	Ito (2007)	
Dvl2	*K14-Dvl2*	Reduced hair follicle length, premature catagen onset	Millar (1999)	
Dkk1	*K14-Dkk1*	Lack of placode formation	Andl (2002)	
	K5rtTA; TRE-Dkk1	Lack of hair follicle neogenesis upon wounding	Ito (2007)	
APC	*K14-Cre; APC $^{fl/fl}$*	Aberrent hair follicle growth, short, curly whiskers	Kuraguchi (2006)	
Lef1 (Transcription factor)	*Lef1 -/-*	Lack of body hair and whisker follicles	van Genderen (1994)	
	K14-Lef1	Clustered hair follicles, irregularly angled shafts	Zhou (1995)	
	K14-ΔNLef1	Suppressed hair follicle differentiation , enhanced sebocyte differentiation	Merrill (2001)	
	K14-ΔNLef1	Progressive hair loss and cyst and tumor formation	Niemann (2002)	
β-catenin	*K14-ΔNβ-catenin (constitutively active)*	*De novo* hair follicle induction	Gat (1998)	
	K14-Cre; β-catenin $^{fl/fl}$	Lack of placode formation, loss of hair in the first hair cycle	Huelsken (2001)	
	K5-S33Yβ-cateninER	Anagen induction and enhanced hair follicle proliferation	Van Mater (2003)	
	K14-ΔNβ-cateninER	Anagen induction and *de novo* hair follicle formation	Lo Celso (2004)	
	K14-ΔNβ-catenin (hemizygous)	Precocious bulge stem cell activation	Lowry (2005)	
	K14-CreER^{tm}; β-catenin $^{fl/fl}$	Loss of bulge stem cell quiescence and maintenance	Lowry (2005)	
	K14-CreER^{tm}; β-catenin $^{fl/fl}$	Lack of hair follicle neogenesis upon wounding	Ito (2007)	
	K14-Cre; Ctnnb1 $^{(ex3)fl/+}$ (Dominant active)	Premature hair follicle placode development, predominant differentiation towards hair shaft, defects in hair follcle development	Zhang (2008)	
	Cor-Cre (DP specific); Ctnnb1 $^{fl/fl}$	Reduced hair follicle length and thinning	Enshell-Seijffers (2010)	
	K15-ΔNβ-cateninER	Increased bulge proliferation and expansion	Baker (2010)	
	ΔK5ΔNβ-cateninER	Ectopic hair follicle induction In sebaceous gland	Baker (2010)	
Tcf3/4 (Transcription factor)	*K14-ΔNTcf3*	Suppressed epidermal differentiation	Merrill (2001)	
	K14rtTA; TRE-mycTcf3	Lack of epidermal differentiation and sebaceous gland	Nguyen (2006)	

Table 2.1. (continued)

2. BMP/TGFβ-related				
TGFβ1	*TGFβ1 -/-*	Advanced hair follicle formation, delayed catagen development	Foitzik (1999), Foitzik (2000)	
TGFβ2	*TGFβ2 -/-*	Delay of hair follicle morphogenesis and reduced number of hair follicles	Foitzik (1999)	
TGFβ3	*TGFβ3 -/-*	No effect	Foitzik (1999)	
TGFβRII	*K14-Cre; TβRII $^{fl/fl}$*	Prolonged telogen length	Oshimori (2011)	
Bmpr1a	*K14-Cre; Bmpr1a $^{fl/fl}$*	Impaired IRS and hair shaft differentiation	Kobielak (2003)	
	K14-CreER^{T2}; Bmpr1a $^{fl/fl}$	Impaired postnatal hair follicle differentiation	Andl (2004)	
	En1-Cre; Bmpr1a $^{fl/fl}$	Lack of hair outgrowth, defects in Matrix cell proliferation and differentiation in morphogenesis and hair cycle	Ming Kwan (2004)	
	Emx1-Cre; Bmpr1a $^{fl/fl}$	Impaired IRS differentiation, reduced number of hair follicle	Yuhki (2004)	
	K14-CreER; Bmpr1a $^{fl/fl}$	Loss of quiescent stem cells	Kobielak (2007)	
Noggin	*Noggin -/-*	Reduction of induced hair placodes	Botchkarev (1999)	
	Msx2-Noggin	Impaired hair shaft differentiation	Kulessa (2000)	
	K14-Noggin	Reduction of telogen length, increased hair density	Plikus (2004), Plikus (2008)	
Activin	*actβA -/-, actβB -/-*	Lack of whisker follicles	Matzuk (1995)	
Follistatin	*fs -/-*	Mis-oriented hair follicles	Matzuk (1995)	

3. Shh-related				
Shh	*Shh -/-*	Arrest in hair follicle morphogenesis after placode formation	St-Jacques (1998), Chiang (1999)	
	K14-Shh	Normal hair placode formation following massive down-growth and tumor formation	Oro (1997)	
	hK14-Shh	Epidermal hyperplasia / loss of epidermal tissue renewal	Adolphe (2004)	
Dhh	*hK14-Dhh*	Epidermal hyperplasia / loss of epidermal tissue renewal	Adolphe (2004)	
Smo	*K5-SmoM2*	Epidermal hyperplasia, but no effect in hair follicle development	Xie (1998)	
	ΔK5-M2SMO	Patchy hair loss and progressive alopecia	Grachtchouk (2003)	
	K14-CreER; RosaSmoM2 / Shh-CreER;RosaSmoM2 / K15-CrePR; RosaSmoM2	BCC originates not from hair follicle bulge but from interfollicular epidermis and the infundibulum	Youssef (2010)	
Gli1,/2/3 (Transcription factor)	*K14-Gli1*	Postnatal hair loss and ulcer development	Oro & Higgins (2003)	
	K5-Gli2	BCC-like tumor formation	Grachtchouk (2000)	

Table 2.1. (continued)

	Gli2-/-	Arrest in hair follicle development	Mill (2003)	
	K5tTA; TRE-Gli2	BCC-like tumors arising from hair follicles	Hutchin (2005)	
Ptch1	*Ptch1^{KO1}/+ or Ptch1^{D11}/+*	Anagen dependent tumor formation	Oro & Higgins (2003)	

4. Notch-related				
Notch	*MHKA1-Notch$^{\Delta E}$ (active form)*	Abnormal differentiation of the medulla and the cuticle	Lin (2000)	
	K5-CreERT; Notch1 $^{fl/fl}$	Epidermal hyperplasia and spontaneous skin tumors	Nicolas (2003)	
	Msx2-Cre; Notch1 $^{fl/fl}$/ Notch2 $^{fl/fl}$/ Notch3 $^{fl/fl}$	Failure to maintain IRS cells, hair follicle cell conversion to epidermal cysts	Pan (2004)	
	Ivl-Notch-IC	Delay in IRS differentiation followed by alopecia	Uyttendaele (2004)	
	K15-CrePR1; Notch1 $^{fl/fl}$/ Notch2 $^{fl/fl}$/ Notch3 $^{-/-}$	Bulge stem cell fate converstion to interfollicular epidermis and formation of epidermal cysts	Demehri (2009)	
RBPj	*Nestin-Cre; RBP-J$^{fl/fl}$*	Hair loss, cyst formation, and epidermal hyperkeratinization	Yamamoto (2003)	
	K14-Cre; RBP-J$^{fl/fl}$	Absense of spinous layer, impaired matrix cell differentiation to IRS and hair shaft	Blanpain (2006)	
	K15-CrePR1; Rbpj$^{fl/fl}$	Bulge stem cell fate conversion to interfollicular epidermis and formation of epidermal cysts	Demehri (2009)	
	K15-CrePR1; Rbpj $^{fl/fl}$ / K5-Cre; Rbpj $^{fl/fl}$	Impaired hair follicle differentiation with mesenchymal RBP-jκ deletion	Hu (2010)	
	Tgfb3-Cre; Rbpj $^{fl/fl}$	Impaired epidermal differentiation	Lin (2011)	
γ-Secretase	*Msx2-Cre; PS1 $^{fl/fl}$ / Msx2-Cre; PS2 $^{fl/fl}$*	Failure to maintain IRS cells, hair follicle cell conversion to epidermal cysts	Pan (2004)	
Jagged1	*K5-Cre; Jag1 $^{fl/fl}$/ K14-CreER; Jag1 $^{fl/fl}$ / K14ΔNβ-cateninER; Jag1 $^{fl/fl}$*	Inhibition of the hair growth cycle and cyst formation	Estrach (2006)	
O-fucosyltransferase 1	*Tgfb3-Cre; Pofut1 $^{fl/fl}$*	Disrupted telogen hair follicle, delay in anagen activation	Lin (2011)	

5. EGF/FGF-related				
TGFα	*Tgfα -/-*	Misoriented, misaligned hair follicles	Luetteke (1993), Mann (1993)	
FGF7	*K14-Fgf7*	Lack of hair follicle formation	Guo (1993)	
FGF5	*Fgf5 -/-*	Elongated hair follicle	Herbert (1994)	
	K5-HERCD-533 (dominant negative)	Short and waved hair follicles and curly whiskers followed by gradual alopecia	Murillas (1995)	
Fgfr2	*Fgfr2-IIIb-/-*	Abnormal hair medullar differentiation	Petiot (2003)	
FGF10	*Fgf10 -/-*	Reduced hair follicle number and abnormal bulb expansion with lack of differerentiated hair shaft	Ohuchi (2000)	

Table 2.1. (continued)

FGF18	*K5-Cre; Fgf18* $^{fl/fl}$	Shortened telogen length	Kimura-Ueki (2012)	

6. Transcription factors	**Mutation type**	**Hair follicle phenotype**	**Reference**	**Related signals**
Runx1	*K14-Cre; Runx1* $^{fl/fl}$ / β-*actinCreER; Runx1* $^{fl/fl}$ / *Runx1* $^{fl/Runx1\text{-}CreER}$	Epithelium: Delayed hair follicle morphogenesis and anagen entry, reduced proliferation; mesenchyme (predominant): abnormal hair follicle morphogenesis and enhanced SG formation	Osorio (2008), Hoi (2010), Osorio (2011)	De-regulates Wnt signaling
Sox9	*Y10-Cre; Sox9* $^{fl/fl}$	Lack of external hair, proliferation defects, ORS cells fate change to epidermal	Vidal (2005)	Downstream of Shh signaling
Lhx2	*K14-Lhx2 (Tg) or Lhx2 -/- (KO)*	Tg; Suppression of hair follicle differentiation; KO: Reduction of hair follicle density but no efffect on differentiation	Rhee (2006)	De-regulates Wnt and Shh downstream factors
NFATc1	*Nfatc1 -/- and K14-Cre; Nfatc1* $^{fl/fl}$	Normal hair follicle morphogenesis, premature stem cell activation and hair follicle growth	Horsley (2008)	Downstrem of Bmp signaling
Tbx1	*K14-Cre; Tbx1* $^{fl/fl}$	Delayed hair cycle entry upon shaving, loss of stem cells upon repeated depilation, reduced proliferative potential	Chen (2012)	Regulates Bmp signaling
Dlx3	*K14-Cre; Dlx3* $^{fl/fl}$	Alopecia due to failure of hair shaft and IRS differentiation	Hwang (2008)	Downstream of Wnt signaling
Msx2 / Foxn1	*Msx2 -/- ; Foxn1 -/-*	Impaired IRS and medulla differentiation	Cai (2009)	De-regulates Notch signaling
p63	*p63 -/- or K5tTA; TRE-ΔNp63α*	Lack of placode formation, defect in hair follicle stem cells and hair cycling	Mills (1999), Yang (1999), Romano (2010)	De-regulates Wnt signaling
Gata3	*Gata3 -/-*	Failure in the IRS precursors differentiation	Kaufman (2003)	Regulates Wnt signaling

ubiquitin-mediated degradation of β-catenin (Clevers and Nusse 2012). Excess of cytoplasmic un-phosphorylated β-catenin molecules translocate into the nucleus, where they interact with Lef/Tcf transcription factors (Vangenderen, Okamura et al. 1994). Wnt transcriptional targets include the Wnt molecules themselves, Eda, Fgfs, Keratins, Lef1, Movo1, Foxn1, Msx2, Follistatin, Tgfβ2, Cyclin D1 and other cell cycle related genes (Clevers and Nusse 2012). Dkk1 is an inhibitor of LRP (Andl, Reddy et al. 2002). β-Catenin also interacts with cadherins (E-cadherin, β-catenin) at the plasma membrane and with APC (Willert and Nusse 1998). Mutation of serine/threonine residues stabilizes β-catenin, mimicking an increased Wnt signal. Recently, LGRs, which are seven transmembrane proteins, were shown to mark stem cell (SC) populations (Jaks, Barker et al. 2008), to physically interact with Frizzled/LRP receptors, and to act in R-spondin-mediated Wnt signaling (de Lau, Barker et al. 2011; Carmon, Lin et al. 2012; Clevers and Nusse 2012).

2.2.2 BMP signaling

Bone Morphogenetic Protein (BMP) signaling consists of two superfamilies of ligands: (a) the BMP family comprising BMPs, GDFs (Growth and Differentiation Factors), and MIS (Muellerian Inhibiting Substance) and (b) the Transforming Growth Factor (TGF)β family comprising TGF I, II, III, Activin, and Nodal. BMP ligands have higher affinity to type I BMP receptors, while TGF family ligands have higher affinity to type II receptors. Type I and II receptors carry an N-terminal ligand-binding domain, a transmembrane domain, and a C-terminal serine/threonine kinase domain (Pelton, Saxena et al. 1991; Shi and Massague 2003). TGFβ/Activin ligands seem to

first bind to type II receptors resulting in a conformational change that enables type I receptors to bind to the complex. Upon activation, receptors can directly phosphorylate R-Smad which releases R-Smad for its binding to Co-Smad (Smad4) and for further localization into the nucleus where it binds the DNA and regulates transcription along with other co-factors (Shi and Massague 2003). R-Smads (Smad 2/3 for TGFβ and Smad 1/5/8 for BMP family) are normally bound to type I receptor in an inactive state. The BMP-related ligands can be bound and inhibited by inhibitors such as Noggin or Follistatin (Schmidt-Ullrich and Paus 2005).

2.2.3 Shh signaling

The mammalian Hedgehog family consists of three secreted proteins: Sonic Hedgehog (Shh), Indian Hedgehog and Desert Hedgehog (Athar, Tang et al. 2006). Patched1, a twelve-span transmembrane protein that acts as a Shh antagonist, normally represses Smoothened (Smo), a seven-span transmembrane co-receptor. Shh signaling inhibits Patched1 thereby relieving the repression of Smo, which can now activate Gli proteins which serve as transcriptional effectors (Callahan and Oro 2001). Each Gli gene (Gli1, 2, and 3) has varying expression patterns in different time and location with differential impact on tissue development and physiology (Matise, Epstein et al. 1998; Matise and Joyner 1999; Park, Bai et al. 2000).

2.2.4 Notch signaling

Notch is a single-span transmembrane protein, originally identified in Drosophila (Kopan and Ilagan 2009), with four family members Notch1, 2, 3, and 4 identified in mammals. The extracellular domain of Notch carries Epidermal Growth Factor (EGF)-like repeats and is activated upon ligand-binding (Delta or Jagged) from neighboring cells. The activation signal triggers ADAM, a metalloproteinase, to cleave Notch before the transmembrane domain and release the intracellular domain from the membrane. The ADAM cleaved product of Notch acts as a substrate for γ-secretase, which cleaves and subsequently releases NICD (Notch Intra-Cellular Domain) for its translocation into the nucleus, where it forms a complex with CBF1/RBPj/Su(H)/Lag-1 for transcriptional activation (Kopan and Ilagan 2009).

2.2.5 Other signaling pathways

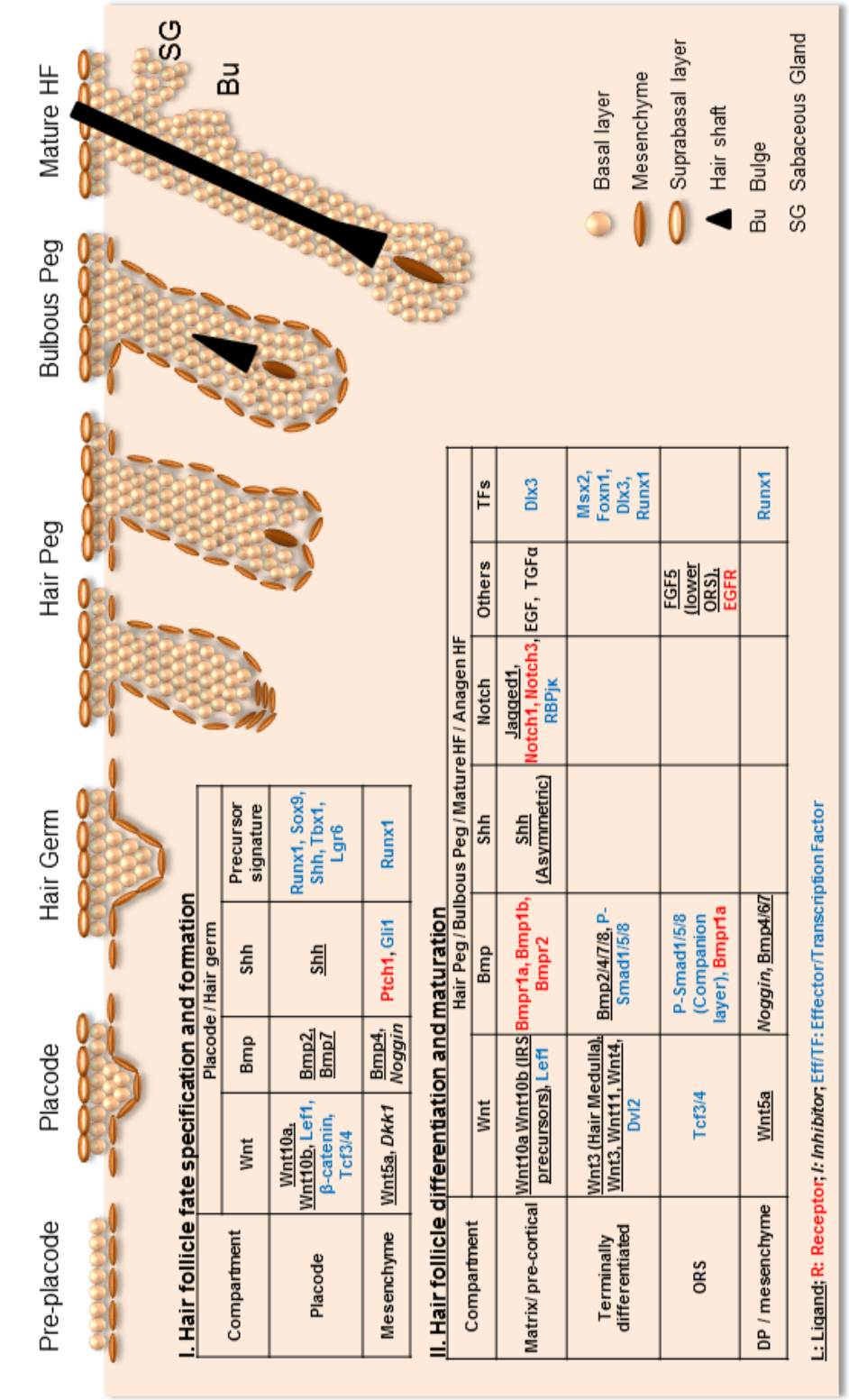
EGFs and TGFβ bind to EGF-Receptors (EGFRs), and this binding triggers dimerization of the receptors and auto-phosphorylation of tyrosine residues, leading to transcriptional activation through Src-Homology proteins, ultimately resulting in cell proliferation (Ullrich and Schlessinger 1990; Pawson and Schlessingert 1993). In particular, EGF molecules promote robust keratinocyte proliferation (Barrandon and Green 1987). Fibroblast Growth Factor (FGF) ligands signal through FGF receptors that are tyrosine-kinases and regulate various cellular processes such as proliferation, migration, and differentiation (Yamasaki, Miyake et al. 1996).

2.2.6 Hair follicle specification, generation, and maintenance during development and skin homeostasis

Skin epithelium arises from the single layer of ectoderm that surrounds the embryo body during embryonic development (Blanpain and Fuchs 2006). Cells of the basal layer (BL) produce highly keratinized terminally differentiated cells of the suprabasal layers that form the interfollicular epidermis. Starting around E14.5, epithelial–mesenchymal signaling triggers BL cells of the epithelium to aggregate, forming a hair placode (Figure 2.1) (Hardy 1992; Tumbar 2012), which proliferate and form a hair germ (HG). The mesenchymal cells residing below start to condense, and as the HG becomes a "peg", they become engulfed in the epithelium forming the "dermal papilla" (DP). Soon after, the highly proliferative cells at the bottom of the peg become matrix (Mx) cells, and start producing inner layers, terminally differentiated HF lineages, forming a bulbous peg (Figure 2.1). Meanwhile, cells on one side of the upper outer root sheath (ORS) cluster, forming a bulge structure that would later contain HFSCs. The sebaceous gland (SG) is also generated from the epithelial HF cells (Frances and Niemann 2012). Melanocytes are a separate lineage that originates from the neural crest, resides in the bulge and Mx, and produces melanin that incorporates into the differentiated hair shaft for color. The HF compartmentalization and differentiation lasts up to ~2.5 weeks after mouse birth (Schmidt-Ullrich and Paus 2005; Blanpain and Fuchs 2006).

After a period of hair growth known as anagen, cells in the lower HF portion undergo apoptosis and HF regresses during the catagen phase, which brings the DP in close proximity to the bulge (Figure 2.2). Around ~PD21 HF enters quiescence or rest

Figure 2.1. Signaling in hair follicle morphogenesis and development



I. Hair follicle fate specification and formation

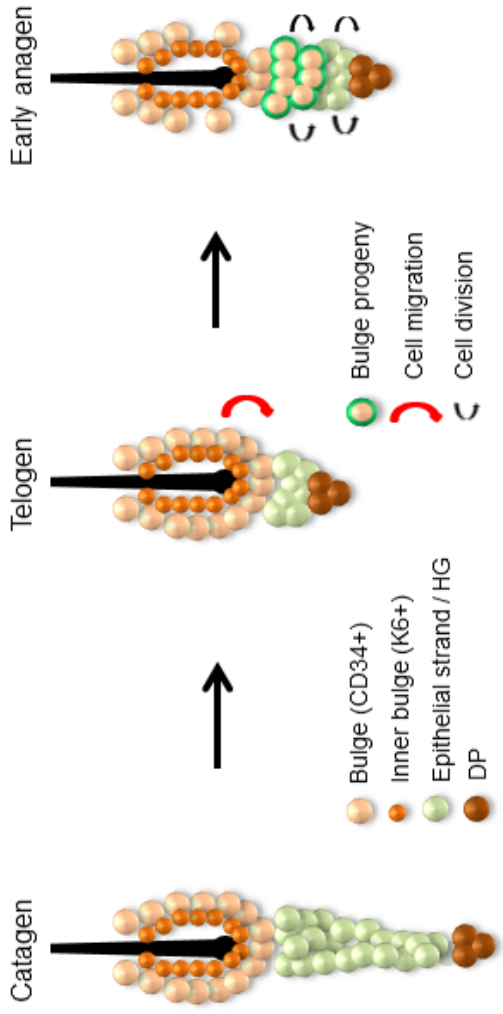
Compartment	Placode / Hair germ			
	Wnt	Bmp	Shh	Precursor signature
Placode	Wnt10a, Wnt10b, Lef1, β-catenin, Tcf3/4	Bmp2, Bmp7	Shh	Runx1, Sox9, Shh, Tbx1, Lgr6
Mesenchyme	Wnt5a, *Dkk1*	Bmp4, *Noggin*	Ptch1, Gli1	Runx1

II. Hair follicle differentiation and maturation

Compartment	Hair Peg / Bulbous Peg / Mature HF / Anagen HF					
	Wnt	Bmp	Shh	Notch	Others	TFs
Matrix/ pre-cortical	Wnt10a Wnt10b (IRS precursors), Lef1	Bmpr1a, Bmp1b, Bmpr2	Shh (Asymmetric)	Jagged1, Notch1, Notch3, RBPjκ	EGF, TGFα	Dlx3
Terminally differentiated	Wnt3 (Hair Medulla), Wnt3, Wnt11, Wnt4, Dvl2	Bmp2/4/7/8, P-Smad1/5/8				Msx2, Foxn1, Dlx3, Runx1
ORS	Tcf3/4	P-Smad1/5/8 (Companion layer), Bmpr1a			FGF5 (lower ORS), EGFR	
DP / mesenchyme	Wnt5a	*Noggin*, Bmp4/6/7				Runx1

L: Ligand; R: Receptor; *I: Inhibitor*; Eff/TF: Effector/Transcription Factor

Figure 2.2. Adult hair follicle signaling from quiescence to activation



Compartment	Refractory telogen				Competent telogen -> Anagen activation			
	Wnt	Bmp	Others	HFSCs signatures	Wnt	Bmp	Others	HFSCs signatures
Bulge (CD34+)	Tcf3/4	Bmp6	Fgf18	Runx1, NFATc1, Lgr5, Sox9, Lhx2	β-catenin, Lef1	Bmp6, TGFβRII, P-Smad2 (LowerBu)	Fgf18	Runx1, NFATc1, Lgr5, Sox9, Lhx2
Hair Germ		Bmp4, Bmpr1a (also in SG)		Runx1	Wnt10a/b, Lef1	TGFβRII, *Noggin*, P-Smad2, Tmeff1		Sox9, Lhx2, Runx1
Inner bulge (K6)		Bmp6	Fgf18			Bmp6	Fgf18	
DP / Mesenchyme		Bmp4			Wnt10a/b	TGFβ2, *Noggin*	Fgf7, Fgf10	
Adipocytes		Bmp2 (Subcutaneous)					PDGFα (Precursors)	

L: Ligand; R: Receptor; *I: Inhibitor*; Eff/TF: Effector/Transcription Factor

(telogen), but is ready to re-activate upon signaling from the DP and the environment. During adult homeostasis, HF undergoes cycles of anagen, catagen and telogen, referred to as "hair cycle" (Cotsarelis 2006). HFSCs reside in the bulge and are the source that fuels all HF lineages (Cotsarelis, Sun et al. 1990; Fuchs, Tumbar et al. 2004; Claudinot, Nicolas et al. 2005). At telogen, upon activation signals, some bulge SCs migrate out into HG where they lose SC characteristics and become progenitors that start to divide rapidly (Zhang, Cheong et al. 2009). The remaining bulge cells divide ~3×during one hair cycle to replenish their SC pool (Waghmare, Bansal et al. 2008; Zhang, White et al. 2010). Rapid proliferation of progenitors (or Transit-Amplifying (TA) cells) results in formation of Mx cells, which give rise to differentiated HF lineages, similar to the anagen of morphogenesis (Muller-Rover, Handjiski et al. 2001). In catagen, cells in the lower HF die again, followed by re-arrangement of the upper and mid ORS cells around the new club hair and the germ, respectively (Hsu, Pasolli et al. 2011). Upon injury conditions such as hair plucking, waxing, and wounding, HFSCs are activated and produce not only HFs but also contribute to epidermal regeneration (Blanpain and Fuchs 2006).

Around the HF, other organs and tissue structures aid in HF homeostasis. The Arrector Pili Muscle (APM) provides the physical force in HF erection necessary for body temperature and appearance (Fujiwara, Ferreira et al. 2011). Recent evidence suggests that Nephronectin, a Wnt-target gene, is enriched in the bulge and this subset of bulge cells serve as a niche for muscle cells. Also, vascularization regulated by VEGF is essential for creating a nurturing environment around the HF (Yano, Brown et al. 2001; Beck, Driessens et al. 2011). In addition, surrounding adipocytes are

important for HF cycling, in which PDGF signaling plays a role (Festa, Fretz et al. 2011). Aberrant signaling is associated with various skin conditions such as pilomatricoma, basal cell carcinomas (BCCs), hamartoma, etc. (Gailani, Stahle-Backdahl et al. 1996; Hahn, Wicking et al. 1996; Dahmane, Lee et al. 1997; Oro, Higgins et al. 1997; Gat, DasGupta et al. 1998; Xie, Murone et al. 1998; Chan, Gat et al. 1999; Grachtchouk, Mo et al. 2000; Nilsson, Unden et al. 2000; Callahan and Oro 2001; Nakamura, Sundberg et al. 2001; Grachtchouk, Grachtchouk et al. 2003; Mill, Mo et al. 2003; Ming Kwan, Li et al. 2004; Yuhki, Yamada et al. 2004), indicating the importance of maintaining the balance of signaling cascades. A comprehensive review underlining the relevance to human disease has been recently published (Boehnke, Falkowska-Hansen et al. 2012). Perhaps the best understood is the aberrations that result from perturbation of the Shh signaling cascade. Components of Shh signaling cascade such as Patched1, Gli1, Smo, and Shh are expressed at high level in many BCCs, and over-expression of individual molecules resulted in BCC-like tumor formation (Gailani, Stahle-Backdahl et al. 1996; Hahn, Wicking et al. 1996; Dahmane, Lee et al. 1997; Oro, Higgins et al. 1997; Xie, Murone et al. 1998; Athar, Tang et al. 2006). Here, we classify these important signaling pathways based on the stage of HF remodeling upon which they act and suggest areas that require further exploration.

2.3 Signaling during hair follicle development

2.3.1 Initiation of hair follicle morphogenesis

Wnt signaling is critical for HF induction (Alonso and Fuchs 2003). Wnt10a, Wnt10b, nuclear -catenin and Lef1 are expressed in the hair placode epithelium (Vangenderen, Okamura et al. 1994; Zhou, Byrne et al. 1995; Merrill, Pasolli et al. 2004), while Wnt5a is expressed in the mesenchyme right below the placode (Reddy, Andl et al. 2001). Targeted deletion of β-catenin in the epithelium via Keratin 14 (K14) promoter active in the BL and ORS inhibits hair placode formation (Huelsken, Vogel et al. 2001), and down-regulation of Lef1, Bmp2, Bmp4, Bmp7, Shh, and Patched1. Similarly, over-expression of Dkk1 results in absence of β-catenin, Wnt10b, Lef1, Edar, Bmp2, Bmp4, Gli1, and K17, and lack of hair placode (Andl, Reddy et al. 2002). Conversely, constitutively active β-catenin induces precocious placode formation and multiple HG budding out from the original follicle (Gat, DasGupta et al. 1998; Silva-Vargas, Lo Celso et al. 2005; Zhang, Andl et al. 2008). β-Catenin and Hedgehog signal strength can specify the number and location of HFs in adult epidermis without recruitment of bulge SCs.

Lef1 skin knockout reduces the HF numbers (Vangenderen, Okamura et al. 1994) and whisker follicle induction, but expression of Lef1 in the dermis at E11.5 rescues this phenotype (Kratochwil, Dull et al. 1996), suggesting a critical role for mesenchymal Lef1. Transgenic Wnt reporter mice containing multiple Lef/Tcf binding sites (TOPGAL) shows activity in the placode as well as in the underlying mesenchyme, which correlates with Lef1 protein localization (DasGupta and Fuchs 1999). Other Wnt-related transcription factors, Tcf3 and 4 are also in the placode, but fade away as HF undergoes further down-growth (Nguyen, Rendl et al. 2006). Double

knockout of Tcf3/4 display loss of hair and thinning of skin (Nguyen, Merrill et al. 2009).

Aside from Wnt-related molecules, the placode expresses Bmp2, Bmp7, and Shh while the underlying mesenchyme expresses Noggin, Bmp4, Patched1, and Gli1 (Reddy, Andl et al. 2001; Andl, Reddy et al. 2002). Deletion of Noggin reduces epithelial placode density except for the guard hairs. A few placodes that form display cytoplasmic localization of β-catenin, which is rescued by Lef1 over-expression (Botchkarev, Botchkareva et al. 1999; Jamora, DasGupta et al. 2003). As epithelial invagination begins, the placode down-regulates E-Cadherin and gains expression of P-Cadherin. E-Cadherin failed to be down-regulated in Noggin knockout mice, and this was also rescued by Lef1 (Jamora, DasGupta et al. 2003).

FGF7 and FGF2 over-expression in K14-lineages also reduce HF density, but display epidermal hyper-proliferation, suggesting that FGF signaling may promote epidermal instead of HF cell fate choice (du Cros 1993; Guo, Yu et al. 1993). While FGF7 null mice does not show a significant phenotype (Guo, Degenstein et al. 1996), over-expression of FGF7 dominant negative form in skin results in defects in HF production and significant reduction in proliferation upon wounding (Werner, Smola et al. 1994). In summary, high level of Wnt signaling and suppression of BMP and FGF signaling are required for epidermal cell fate choice toward HF cells and allow these cells to aggregate and begin to invaginate to form the hair placode.

2.3.2 Follicular invagination and maturation

Shh knockout or knockdown show initial epithelium invagination and mesenchymal condensation, but HFs become developmentally arrested at this stage (St-Jacques, Dassule et al. 1998; Chiang, Swan et al. 1999; Wang, Liu et al. 2000). Shh transcriptional targets (Ptc1, Smo, Gli1 and Gli2) are detected at lower levels, but Wnt and Bmp signaling are unaffected by the Shh knockout, suggesting that Shh signaling is not critical for initial placode formation but is important for HF growth. Gli2 knockout mice show similar phenotype to the Shh knockout mice (Mill, Mo et al. 2003). K14-ΔNβ-catenin mice show aberrant expression pattern of Shh, often on both sides of Mx, and this is associated with wavy HFs, suggesting that asymmetric Shh localization in the Mx is critical for follicular polarity (Gat, DasGupta et al. 1998).

Wnt signaling plays important roles during follicular growth. For instance, Wnt3 expression is confined to differentiating hair medulla (Millar, Willert et al. 1999). Dvl2, another important player of the Wnt pathway, co-localizes with AE13+ differentiated cortex layer. K14-Wnt3 and K14-Dvl2 transgenesis result in shortening of hair shafts. In addition, Wnt10a/Wnt10b are in the IRS precursors while Wnt5a is expressed in the DP and a subset of ORS cells. Wnt11, Wnt3a, and Wnt4 are present in specific HF layers during maturation (Reddy, Andl et al. 2001). Nuclear Lef1 is confined to the tip of the growing follicle and the mesenchyme touching the tip. At birth, Lef1 mRNA is found in the Mx and pre-cortical region, and Lef1 protein nuclear localization becomes strong in the pre-cortex (DasGupta and Fuchs 1999). Interestingly, TOPGAL mice show significant signal enrichment in the hair cortex. On the other hand, Tcf3 is not expressed in the Mx but in the lower ORS.

Targeted deletion of β-catenin in melanocyte SCs results in loss of hair pigmentation, along with reduced proliferation and absence of differentiation markers (Rabbani, Takeo et al. 2011). Interestingly, induction of β-catenin dominant active form in K15 lineage results in expansion of melanocytes, promoting their proliferation. This seemed to be regulated through Edn, a direct Wnt target gene that is highly induced upon β-catenin activation in epithelial cells, and concomitant increased expression of EdnrB in melanocyte SCs.

Bmp signaling is critical for proper HF differentiation. Bmpr1a deletion results in abnormal HF formation due to reduced p-Smad1/5/8 levels through which Bmp signaling occurs along with a lack of lineage-specific differentiation markers (Andl, Ahn et al. 2004).

In conjunction with common signaling pathways, different hair types (vibrissae, guard, and pelage) seem to employ slightly different signals for HF induction. For instance, EdaA1/EdaR signaling through NF-κB is essential for guard hair specification (Headon and Overbeek 1999; Mikkola, Pispa et al. 1999; Schmidt-Ullrich, Aebischer et al. 2001). On the other hand, mutations in ActivinA or its inhibitor Follistatin results in defects in development of vibrissae and pelage HFs (Munz, Smola et al. 1999; Laurikkala, Pispa et al. 2002; Nakamura, Matzuk et al. 2003). Sox2 is expressed in DP of "some" but not all HF types and is important for the specification of these follicular types (Driskell, Giangreco et al. 2009).

While Mx cells give rise to differentiated HF lineages, one side of upper ORS cells generate a cluster, forming the bulge, which contains adult HFSCs that are essential for adult HF cycling and skin homeostasis. These differential fates seemed to

be specified early in the placode stage, where there is heterogeneity of several molecular markers such as Shh (Levy, Lindon et al. 2005), Sox9 (Nowak, Polak et al. 2008), Lgr6 (Snippert, Haegebarth et al. 2010), Tbx1 (Chen, Heller et al. 2012) and Runx1 (Osorio, Lilja et al. 2011) expression in all or a subset of cells in the placode. Lineage tracing experiments showed that these placode cells give rise to ORS, bulge, and other HF lineages (Tumbar 2012). More importantly, these factors interact with signaling molecules in regulating proper HF development. For instance, Runx1 expression is detected in a subset of placode cells as well as in mesenchyme in the very early stage of morphogenesis. Runx1 expression becomes very prominent through the growing follicle as well as in DP, and interacts with Lef1 in differential mediation of Wnt signaling for proper emergence and maintenance of HFSCs (Osorio, Lilja et al. 2011). Runx1 expressing cells eventually give rise to adult HFSCs, and Runx1 plays additional roles in their proliferation in adulthood (Osorio, Lee et al. 2008; Hoi, Lee et al. 2010). Mature HFs are compartmentalized and can be characterized by specific expression profiles. The Gli1+ region above the bulge and the Lrig1+, Mts24+, Lgr6+, and Blimp1+ region of the isthmus and SG bear SCs, but these do not contribute to long-term HF regeneration (Horsley, O'Carroll et al. 2006; Nijhof, Braun et al. 2006; Jensen, Collins et al. 2009; Snippert, Haegebarth et al. 2010; Brownell, Guevara et al. 2011).

In addition to Wnt, Bmp, and Shh signaling, Notch signaling also ensures proper HF differentiation. Notch1 is expressed in the Mx. Notch2 is not detected, but Notch3 and Jagged1 are expressed in pre-cortex cells and cuticle layer of IRS (Watt, Estrach et al. 2008). Over-expression of Notch1 Intra-Cellular Domain (NICD) in pre-

cortical and cortical cells results in hair loss at the end of 1st hair cycle and displays abnormal differentiation (Lin, Leimeister et al. 2000; Uyttendaele, Panteleyev et al. 2004). Loss of Notch1/2 as well as loss of RBPjκ or γ-secretase in epithelium, result in initiation of HF morphogenesis, but show similar defects of abnormal differentiation followed by hair loss (Pan, Lin et al. 2004; Blanpain, Lowry et al. 2006; Cai, Lee et al. 2009; Demehri and Kopan 2009). RBPjκ deletion in DP (ColICre) also shows disrupted IRS followed by HF degeneration and cyst formation (Hu, Lefort et al. 2010). In addition, Wnt5a, Fgf7, Fgf10 and Noggin are down-regulated in the DP and re-introduction of these factors partially rescued the HF degeneration phenotype. These results suggest that Notch1 signaling controls DP signature molecules that in return signal to the Mx cells to differentiate.

Importantly, Notch signaling works downstream of several pathways. For instance, ectopic expression of constitutively active β-catenin in K14-lineages results in up-regulation of Notch pathway genes such as Jagged1, suggesting a Wnt-dependent regulation of Notch signaling (Estrach, Ambler et al. 2006; Ambler and Watt 2007; Watt, Estrach et al. 2008). Msx2 and Foxn1 seem to play redundant roles in regulating Notch1 expression and are important for proper HF differentiation. Interestingly, Bmpr1a deficient mice display reduction in Msx2 and Foxn1 expression in the HF and show similar phenotype to Msx2/Foxn1 double mutant, suggesting that BMP signaling from Mx/DP regulate Msx2/Foxn1-mediated Notch1 expression in proper HF differentiation (Andl, Ahn et al. 2004; Cai, Lee et al. 2009). Finally, EGF/TGFβ molecules are also detected in the hair bulb and IRS and in the absence of

TGFβ, HFs become wavy and curly, displaying an abnormal differentiation (Luetteke, Qiu et al. 1993; Mann, Fowler et al. 1993).

2.3.3 Regression and quiescence of mature hair follicle

Around ~PD16–17, the lower portion of the mature HF is destructed by apoptosis (catagen). DP moves upward and comes in close contact to the bulge, and the HF enters telogen phase around ~PD20–21. Interestingly, targeted deletion of β-catenin in the DP (Corin-Cre) around PD3 during HF growth phase, results in premature entrance into catagen (Enshell-Seijffers, Lindon et al. 2010). Dlx3, a Dlx-related transcription factor, regulates timely HF regression from the epithelial compartment (Hwang, Mehrani et al. 2008). Dlx3 expression extends from Mx to IRS, and Lef1 regulates its expression. Dlx3 mutant HFs show enhanced proliferation along with delayed regression due to the absence of BMP signaling (Hwang, Mehrani et al. 2008). FGF5 is expressed in late anagen during morphogenesis, and its deletion delays catagen, resulting in longer hair shafts (Enshell-Seijffers, Lindon et al. 2010).

In summary, multiple signals including Wnt, Bmp, Shh, Notch, FGF and EGF are at work in proper follicular morphogenesis: growth, lineage specification, and timely regression after maturation. Interestingly, a handful of the same signaling pathways are utilized in both formation and maturation of the HFs. However, the different combinations of these pathways during a specific time window and cell populations likely program the cells to distinct fates.

2.4 Signaling in hair follicle cycling during homeostasis and injury

2.4.1 From quiescence to activation

The adult hair cycle begins at first telogen (~PD20–21) after HFs finished regression and await activation (Figure 2.2). HF quiescence is maintained via Bmp signaling (Rendl, Lewis et al. 2005). Specifically, Bmp4 is expressed in DP and secondary HG while Noggin, a Bmp inhibitor, is absent in telogen follicle. On the other hand, Bmpr1a is expressed in HG and in SG (Botchkarev, Botchkareva et al. 2001). In addition, Bmp2/4 levels are highest at telogen. Inhibition of Bmp signaling via Noggin over-expression resulted in precocious anagen entrance upon hair plucking, suggesting a dominant role of Bmp signaling in maintaining quiescence (Plikus, Mayer et al. 2008). HFSCs quiescence is also maintained via Bmp regulation of NFATc1, a bulge-specific marker that once lost results in increased proliferation (Horsley, Aliprantis et al. 2008). A recent study also suggests that TGFβ signaling is a major pathway regulating the adult hair cycle (Oshimori and Fuchs 2012). TGFβ receptor II (TGFβRII) activity and p-Smad2 were detected in the lower bulge and the HG at the transition from telogen to anagen. In addition, late-telogen DP is enriched with TGFβ2 ligand and injection of TGFβ2 ligand in the TGFβRII knockout mice before endogenous TGFβ2 expression in DP failed to activate anagen at a normal time due to a direct regulation of Tmeff1, a transmembrane protein that inhibits BMP signaling (Oshimori and Fuchs 2012). Moreover, the inner layer (K6+) of the bulge expresses significant amounts of Bmp6 and Fgf18, which inhibit the outer bulge stem cell layer (CD34+) proliferation (Figure 2.2). Fgf18 deletion in Keratin 5 (K5)+ cells (BL and ORS) results in shorter telogen length before anagen re-entry. In addition,

subcutaneous Fgf18 as well as BMP and Wnt inhibitors all induce HF quiescence (Plikus, Baker et al. 2011; Kimura-Ueki, Oda et al. 2012). Finally, ablation of the inner bulge layer (K6+) induces bulge cell proliferation, suggesting a paracrine role of Bmp6 and Fgf18 in maintaining bulge SC quiescence (Hsu, Pasolli et al. 2011).

Upon activation, Bmp4 expression in DP and HG decreases while the rest of the skin dermis shows Bmp4 up-regulation (Kobielak, Stokes et al. 2007). Concomitantly, DP and the growing HG up-regulate Noggin and down-regulate Bmpr1a. Deletion of Bmpr1a in telogen following shaving results in precocious bulge cells proliferation, but no further development of HF. Concomitantly there is loss of CD34 in the bulge as well as Sox9 and Lhx2 expression in the HG (Kobielak, Stokes et al. 2007). In addition, the bulge shows increased nuclear β-catenin and Lef1, which is mediated by inactivation of GSK-3β by its increased phosphorylation, potentially driven by PTEN/PI3-K/P-Akt pathway.

Low level of Wnt signaling is also essential for maintaining HF quiescence. K14-ΔNβ-catenin hemizygous mice display precocious anagen activation, revealing a dose-dependence of β-catenin in hair cycle activation (Lowry, Blanpain et al. 2005). In addition, inducible ectopic expression of activated β-catenin in the 2nd telogen results in precocious induction of anagen-like HF structure associating hyper-proliferation as well as induction of *de novo* sites of HF formation (Van Mater, Kolligs et al. 2003; Lo Celso, Prowse et al. 2004). Lef1 is absent in the bulge, but is enriched in HG during telogen. On the other hand, Tcf3/4 are highly enriched in the bulge, originating from the embryonic Tcf3/4 expressing cells of the ORS (Nguyen, Rendl et al. 2006), acting as a repressor of Wnt signaling (Merrill, Gat et al. 2001; Nguyen,

Merrill et al. 2009). Conversely, Wnt10a/Wnt10b are enriched in the bottom of the HG and in DP at telogen–anagen transition (Reddy, Andl et al. 2001). FGF7 expression gradually increases in DP during the long 2nd telogen prior to activation. Moreover, ectopic Fgf7 induction in telogen results in proliferation induction in the HG (Greco, Chen et al. 2009). Some of the Wnt pathway genes may be regulated by Runx1, which is present in the lower bulge and HG and governs timely HFSCs activation and proliferation (Osorio, Lee et al. 2008; Hoi, Lee et al. 2010).

Lgr5, a protein that is enriched in a subset of cells in the lower bulge and the HG in telogen bear multipotency and seem to have Wnt-dependent function. *In vitro* and biochemical assays revealed that LGR proteins bind to Frizzled and LRP receptors forming a complex, which can activate canonical Wnt signaling pathway (Jaks, Barker et al. 2008; de Lau, Barker et al. 2011; Carmon, Lin et al. 2012). The role of these LGR proteins in HF *in vivo* has not yet been addressed.

Recently, a circadian molecular clock mediated via Bmal1/Clock transcription factors was shown to affect HFSCs proliferation and quiescence through de-regulation of many critical signaling molecules such as Wnt-related (Dkk3, Lef1, Fzd2/3, Sox9, Lhx2, Lgr5, and Tcf3) and Bmp-related (Smad7, Ltbps, and Lefty) (Janich, Pascual et al. 2011). Wnt, Bmp, and Fgf interactions counterbalance each other inhibitory and activation signals and predispose telogen follicles to be at a “refractory” or “competent” status with respect to anagen entry (Plikus 2012). The “waves” of some of these signaling molecules traveling as gradients throughout the mouse skin across the body may be the consequence of fluctuations of the components in the circadian clock that regulate signaling pathway genes.

2.4.2 Growth and differentiation of adult hair follicle

Growth and differentiation of adult HF employs similar signaling as morphogenesis. Mx cells and pre-cortex cells express Wnt3 and Wnt10a/b (Reddy, Andl et al. 2001). These molecules send cascades of signaling to Lef1, which also show enrichment in the Mx and pre-cortical cells (DasGupta and Fuchs 1999). In addition, differentiated layers express many Wnt molecules such as Wnt3a, Wnt4, and Wnt11.

In addition to Wnt signals, a variety of Bmp ligands and receptors are expressed in the Mx cells (Bmpr1a/b and Bmpr2), hair shaft precursors (Bmp2/4), and IRS lineages (Bmp7/8), which allow the TA cells to stop proliferation and enter the terminal differentiation program (Rendl, Lewis et al. 2005). Moreover, surrounding the HF, Bmp2/4 are present at low level in the mesenchyme during early anagen. These levels gradually peak at the late anagen and telogen to inhibit proliferation (Plikus, Mayer et al. 2008).

The extent of bulge SC proliferation during anagen seems to inversely correlate with the level of BMP ligands (Waghmare, Bansal et al. 2008; Zhang, Cheong et al. 2009). During anagen, bulge SCs self-renew to make up for the loss that occurred prior to anagen activation (Zhang, Cheong et al. 2009). Extended self-renewal push cells out downward into the ORS region below the bulge, which later move upwards to make a new bulge and HG (Hsu, Pasolli et al. 2011). Although to a lesser extent than the inner bulge layer (K6+), bulge SCs (CD34+) express Bmp6 and

Fgf18 that seems to send the inhibitory signal toward the ORS cells in regulating their proliferation during anagen (Blanpain, Lowry et al. 2004; Greco, Chen et al. 2009).

Notch signaling is also important for proper adult HF differentiation, as it was for the developing HF during morphogenesis. RBPjκ targeted deletion in the bulge (K15+) at 2nd telogen results in normal HF activation and growth, but HFs display abnormal terminal differentiation, and in long-term form cysts containing epidermal lineage keratinocytes (Demehri and Kopan 2009).

2.4.3 Destruction and returning back to quiescence in adulthood

Once the proliferative signal fades and inhibitory signals accumulate in the anagen–catagen transition, HF starts to regress from the bottom. TGFβ1 expression is seen in bulge and ORS in early anagen, followed by enrichment in ORS in late anagen (Foitzik, Lindner et al. 2000). Its expression increases in catagen in the bulge and the epithelial strands, and disappears in telogen. TGFβ1 deletion resulted in delayed catagen entrance, and ectopic TGFβ1 expression induced premature catagen, suggesting its importance in timing the switch from proliferation to regression.

BMP signals start to accumulate from late anagen, last throughout catagen, and peak at telogen (Plikus, Mayer et al. 2008). Also, Fgf18 starts to be expressed in catagen and is maintained throughout telogen. Ectopic induction of Fgf18 in anagen shows reduction in proliferation, suggesting its up-regulation during hair cycle may be important for inducing catagen (Kimura-Ueki, Oda et al. 2012). EGFR expression is detected in the ORS (Luetteke, Phillips et al. 1994). Dominant negative EGFR showed initial HF formation, but failed to enter catagen and resulted in hair loss over time

(Murillas, Larcher et al. 1995). Other signaling molecules such as neurotrophins have also revealed their importance in proper regression (Botchkarev, Welker et al. 1998; Botchkarev, Botchkareva et al. 1999; Botchkareva, Botchkarev et al. 2000; Botchkarev, Botchkareva et al. 2001).

In summary, Bmp-mediated repression dominantly contributes to adult HF and HFSCs quiescence, and accumulation of Wnt or other signals are necessary to overcome and lead cells into proliferation. HF bulge cells display extra quiescence relative to other HF compartments due to autocrine and paracrine Bmp and Fgf signals that are innate (from the CD34+ bulge cells) and secreted from the surrounding mesenchyme and the inner bulge layer (K6+). Wnt, Bmp, Shh, Notch, and Egf signals are important for proper adult HF growth and differentiation program. Finally, gradual buildup of Bmp and Fgf signals overtake the proliferative signals, induces HF regression followed by quiescence (telogen).

2.4.4 The curious case of Runx1: when timing and localization matters

As described so far, the growth and differentiation of HF during the morphogenesis and hair cycling resemble each other in their cellular and molecular processes. However, a major difference is that while secondary (adult) HG consists of short-lived progenitors, the primary (embryonic) HG contains precursors that would later give rise to long-lived HFSCs. Interestingly, there is strong and transient expression of Runx1 in the embryonic dermis, which molecularly distinguishes embryonic from adult skin. This coincides with another unique, non-cyclic, event in the life of the HF: the birth of adult HFSCs (Osorio, Lilja et al. 2011). When Runx1 is

knocked out in the mesenchyme during this stage, HF form and go through the first catagen but adult HFSC cells convert their fates toward SG later on in life. Strikingly, Runx1 regulates autocrine and paracrine Wnt signaling in opposing direction in the dermis versus epidermis, thereby creating gradients of signaling in the embryonic skin. Thus, loss of Runx1 in the mesenchyme at the embryonic stage revealed the importance of Wnt signaling in the embryonic environment to nurture the nascent adult HFSCs to mature well, to shape up their molecular makeup and to endow them with long life and multipotency.

2.5 Conclusions and future perspectives

In this review, we summarized our current understandings of signaling pathways that play essential roles in governing the development and cycling of HFs. Mouse genetics allowed researchers to discover the tip of the iceberg in the molecular networks that are participating in the greater purpose of tissue development and homeostasis. Perturbation in these pathways leads to disease and some of the signals unveiled here are already making their useful way to the clinic. As with most human disease, those affecting skin are likely subtly affecting multiple genes and are difficult to model by gene knockout. The generation of induced pluripotent SCs (iPS) (Takahashi and Yamanaka 2006) and the ability to differentiate them to keratinocyte (Hanna, Wernig et al. 2007) combined with high throughput sequencing might allow us not only to identify those gene variations that are possibly associated with disease, but also to model these diseases by regenerating skin and hair follicles from human

iPS in immune-compromised mice. Advances in transgenic mouse technology allow us to relatively easily study single gene mutations (Dow and Lowe 2012) and may in the future be adapted to rapidly create multiple mutations that can mimic specific traits discovered in human disease.

The advances within the field have begun to uncover the veil of skin biology, pointing the researchers to the areas of further exploration. How are signals sent from dermis, adipocytes, blood vessels, and inner bulge cells regulated? To what degree do paracrine and autocrine signals that surround or are innate influence the HF status? Are there more signaling pathways waiting to be discovered? What is the degree of redundancy across different activation or inhibitory signals? How can these signaling networks be used or manipulated in medical purpose of treating patients with skin diseases or wounds? Answers to these questions will bring insights in discovering what remains below the tip of the iceberg of skin biology.

REFERENCES

Alonso, L. and E. Fuchs (2003). "Stem cells in the skin: waste not, Wnt not." Genes Dev **17**(10): 1189-1200.

Ambler, C. A. and F. M. Watt (2007). "Expression of Notch pathway genes in mammalian epidermis and modulation by beta-catenin." Dev Dyn **236**(6): 1595-1601.

Andl, T., K. Ahn, et al. (2004). "Epithelial Bmpr1a regulates differentiation and proliferation in postnatal hair follicles and is essential for tooth development." Development **131**(10): 2257-2268.

Andl, T., S. T. Reddy, et al. (2002). "WNT signals are required for the initiation of hair follicle development." Dev Cell **2**(5): 643-653.

Athar, M., X. Tang, et al. (2006). "Hedgehog signalling in skin development and cancer." Exp Dermatol **15**(9): 667-677.

Barrandon, Y. and H. Green (1987). "Cell migration is essential for sustained growth of keratinocyte colonies: the roles of transforming growth factor-alpha and epidermal growth factor." Cell **50**(7): 1131-1137.

Beck, B., G. Driessens, et al. (2011). "A vascular niche and a VEGF-Nrp1 loop regulate the initiation and stemness of skin tumours." Nature **478**(7369): 399-403.

Blanpain, C. and E. Fuchs (2006). "Epidermal stem cells of the skin." Annu Rev Cell Dev Biol **22**: 339-373.

Blanpain, C., W. E. Lowry, et al. (2004). "Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche." Cell **118**(5): 635-648.

Blanpain, C., W. E. Lowry, et al. (2006). "Canonical notch signaling functions as a commitment switch in the epidermal lineage." Genes Dev **20**(21): 3022-3035.

Boehnke, K., B. Falkowska-Hansen, et al. (2012). "Stem cells of the human epidermis and their niche: composition and function in epidermal regeneration and carcinogenesis." Carcinogenesis.

Botchkarev, V. A., N. V. Botchkareva, et al. (2001). "Noggin is required for induction of the hair follicle growth phase in postnatal skin." FASEB J **15**(12): 2205-2214.

Botchkarev, V. A., N. V. Botchkareva, et al. (1999). "Noggin is a mesenchymally derived stimulator of hair-follicle induction." Nat Cell Biol **1**(3): 158-164.

Botchkarev, V. A., N. V. Botchkareva, et al. (1999). "A new role for neurotrophins: involvement of brain-derived neurotrophic factor and neurotrophin-4 in hair cycle control." FASEB J **13**(2): 395-410.

Botchkarev, V. A., P. Welker, et al. (1998). "A new role for neurotrophin-3: involvement in the regulation of hair follicle regression (catagen)." Am J Pathol **153**(3): 785-799.

Botchkareva, N. V., V. A. Botchkarev, et al. (2000). "New roles for glial cell line-derived neurotrophic factor and neurturin: involvement in hair cycle control." Am J Pathol **156**(3): 1041-1053.

Brownell, I., E. Guevara, et al. (2011). "Nerve-derived sonic hedgehog defines a niche for hair follicle stem cells capable of becoming epidermal stem cells." Cell Stem Cell **8**(5): 552-565.

Cai, J., J. Lee, et al. (2009). "Genetic interplays between Msx2 and Foxn1 are required for Notch1 expression and hair shaft differentiation." Dev Biol **326**(2): 420-430.

Callahan, C. A. and A. E. Oro (2001). "Monstrous attempts at adnexogenesis: regulating hair follicle progenitors through Sonic hedgehog signaling." Curr Opin Genet Dev **11**(5): 541-546.

Carmon, K. S., Q. Lin, et al. (2012). "LGR5 Interacts and Cointernalizes with Wnt Receptors To Modulate Wnt/beta-Catenin Signaling." Mol Cell Biol **32**(11): 2054-2064.

Chan, E. F., U. Gat, et al. (1999). "A common human skin tumour is caused by activating mutations in beta-catenin." Nat Genet **21**(4): 410-413.

Chen, T., E. Heller, et al. (2012). "An RNA interference screen uncovers a new molecule in stem cell self-renewal and long-term regeneration." Nature **485**(7396): 104-108.

Chiang, C., R. Z. Swan, et al. (1999). "Essential role for Sonic hedgehog during hair follicle morphogenesis." Dev Biol **205**(1): 1-9.

Claudinot, S., M. Nicolas, et al. (2005). "Long-term renewal of hair follicles from clonogenic multipotent stem cells." Proc Natl Acad Sci U S A **102**(41): 14677-14682.

Clevers, H. and R. Nusse (2012). "Wnt/beta-Catenin Signaling and Disease." Cell **149**(6): 1192-1205.

Cotsarelis, G. (2006). "Epithelial stem cells: a folliculocentric view." J Invest Dermatol **126**(7): 1459-1468.

Cotsarelis, G., T. T. Sun, et al. (1990). "Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis." Cell **61**(7): 1329-1337.

Dahmane, N., J. Lee, et al. (1997). "Activation of the transcription factor Gli1 and the Sonic hedgehog signalling pathway in skin tumours." Nature **389**(6653): 876-881.

DasGupta, R. and E. Fuchs (1999). "Multiple roles for activated LEF/TCF transcription complexes during hair follicle development and differentiation." Development **126**(20): 4557-4568.

de Lau, W., N. Barker, et al. (2011). "Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling." Nature **476**(7360): 293-297.

Demehri, S. and R. Kopan (2009). "Notch signaling in bulge stem cells is not required for selection of hair follicle fate." Development **136**(6): 891-896.

Dow, Lukas E. and Scott W. Lowe (2012). "Life in the Fast Lane: Mammalian Disease Models in the Genomics Era." Cell **148**(6): 1099-1109.

Driskell, R. R., A. Giangreco, et al. (2009). "Sox2-positive dermal papilla cells specify hair follicle type in mammalian epidermis." Development **136**(16): 2815-2823.

du Cros, D. L. (1993). "Fibroblast growth factor influences the development and cycling of murine hair follicles." Dev Biol **156**(2): 444-453.

Enshell-Seijffers, D., C. Lindon, et al. (2010). "beta-catenin activity in the dermal papilla regulates morphogenesis and regeneration of hair." Dev Cell **18**(4): 633-642.

Estrach, S., C. A. Ambler, et al. (2006). "Jagged 1 is a beta-catenin target gene required for ectopic hair follicle formation in adult epidermis." Development **133**(22): 4427-4438.

Festa, E., J. Fretz, et al. (2011). "Adipocyte lineage cells contribute to the skin stem cell niche to drive hair cycling." Cell **146**(5): 761-771.

Foitzik, K., G. Lindner, et al. (2000). "Control of murine hair follicle regression (catagen) by TGF-beta1 in vivo." FASEB J **14**(5): 752-760.

Frances, D. and C. Niemann (2012). "Stem cell dynamics in sebaceous gland morphogenesis in mouse skin." Dev Biol **363**(1): 138-146.

Fuchs, E., T. Tumbar, et al. (2004). "Socializing with the neighbors: stem cells and their niche." Cell **116**(6): 769-778.

Fujiwara, H., M. Ferreira, et al. (2011). "The basement membrane of hair follicle stem cells is a muscle cell niche." Cell **144**(4): 577-589.

Gailani, M. R., M. Stahle-Backdahl, et al. (1996). "The role of the human homologue of Drosophila patched in sporadic basal cell carcinomas." Nat Genet **14**(1): 78-81.

Gat, U., R. DasGupta, et al. (1998). "De Novo hair follicle morphogenesis and hair tumors in mice expressing a truncated beta-catenin in skin." Cell **95**(5): 605-614.

Grachtchouk, M., R. Mo, et al. (2000). "Basal cell carcinomas in mice overexpressing Gli2 in skin." Nat Genet **24**(3): 216-217.

Grachtchouk, V., M. Grachtchouk, et al. (2003). "The magnitude of hedgehog signaling activity defines skin tumor phenotype." EMBO J **22**(11): 2741-2751.

Greco, V., T. Chen, et al. (2009). "A two-step mechanism for stem cell activation during hair regeneration." Cell Stem Cell **4**(2): 155-169.

Gumbiner, B. M. (1997). "Carcinogenesis: a balance between beta-catenin and APC." Curr Biol **7**(7): R443-446.

Guo, L., L. Degenstein, et al. (1996). "Keratinocyte growth factor is required for hair development but not for wound healing." Genes Dev **10**(2): 165-175.

Guo, L., Q. C. Yu, et al. (1993). "Targeting expression of keratinocyte growth factor to keratinocytes elicits striking changes in epithelial differentiation in transgenic mice." EMBO J **12**(3): 973-986.

Hahn, H., C. Wicking, et al. (1996). "Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome." Cell **85**(6): 841-851.

Hanna, J., M. Wernig, et al. (2007). "Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin." Science **318**(5858): 1920-1923.

Hardy, M. H. (1992). "The secret life of the hair follicle." Trends Genet **8**(2): 55-61.

Headon, D. J. and P. A. Overbeek (1999). "Involvement of a novel Tnf receptor homologue in hair follicle induction." Nat Genet **22**(4): 370-374.

Hoi, C. S., S. E. Lee, et al. (2010). "Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin." Mol Cell Biol **30**(10): 2518-2536.

Horsley, V., A. O. Aliprantis, et al. (2008). "NFATc1 balances quiescence and proliferation of skin stem cells." Cell **132**(2): 299-310.

Horsley, V., D. O'Carroll, et al. (2006). "Blimp1 defines a progenitor population that governs cellular input to the sebaceous gland." Cell **126**(3): 597-609.

Hsu, Y. C. and E. Fuchs (2012). "A family business: stem cell progeny join the niche to regulate homeostasis." Nat Rev Mol Cell Biol **13**(2): 103-114.

Hsu, Y. C., H. A. Pasolli, et al. (2011). "Dynamics between stem cells, niche, and progeny in the hair follicle." Cell **144**(1): 92-105.

Hu, B., K. Lefort, et al. (2010). "Control of hair follicle cell fate by underlying mesenchyme through a CSL-Wnt5a-FoxN1 regulatory axis." Genes Dev **24**(14): 1519-1532.

Huelsken, J., R. Vogel, et al. (2001). "beta-Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin." Cell **105**(4): 533-545.

Hwang, J., T. Mehrani, et al. (2008). "Dlx3 is a crucial regulator of hair follicle differentiation and cycling." Development **135**(18): 3149-3159.

Jaks, V., N. Barker, et al. (2008). "Lgr5 marks cycling, yet long-lived, hair follicle stem cells." Nat Genet **40**(11): 1291-1299.

Jamora, C., R. DasGupta, et al. (2003). "Links between signal transduction, transcription and adhesion in epithelial bud development." Nature **422**(6929): 317-322.

Janich, P., G. Pascual, et al. (2011). "The circadian molecular clock creates epidermal stem cell heterogeneity." Nature **480**(7376): 209-214.

Jensen, K. B., C. A. Collins, et al. (2009). "Lrig1 expression defines a distinct multipotent stem cell population in mammalian epidermis." Cell Stem Cell **4**(5): 427-439.

Kaufman, C. K., P. Zhou, et al. (2003). "GATA-3: an unexpected regulator of cell lineage determination in skin." Genes Dev **17**(17): 2108-2122.

Kimura-Ueki, M., Y. Oda, et al. (2012). "Hair cycle resting phase is regulated by cyclic epithelial FGF18 signaling." J Invest Dermatol **132**(5): 1338-1345.

Kobielak, K., N. Stokes, et al. (2007). "Loss of a quiescent niche but not follicle stem cells in the absence of bone morphogenetic protein signaling." Proc Natl Acad Sci U S A **104**(24): 10063-10068.

Kopan, R. and M. X. Ilagan (2009). "The canonical Notch signaling pathway: unfolding the activation mechanism." Cell **137**(2): 216-233.

Kratochwil, K., M. Dull, et al. (1996). "Lef1 expression is activated by BMP-4 and regulates inductive tissue interactions in tooth and hair development." Genes Dev **10**(11): 1382-1394.

Laurikkala, J., J. Pispa, et al. (2002). "Regulation of hair follicle development by the TNF signal ectodysplasin and its receptor Edar." Development **129**(10): 2541-2553.

Levy, V., C. Lindon, et al. (2005). "Distinct stem cell populations regenerate the follicle and interfollicular epidermis." Dev Cell **9**(6): 855-861.

Lin, M. H., C. Leimeister, et al. (2000). "Activation of the Notch pathway in the hair cortex leads to aberrant differentiation of the adjacent hair-shaft layers." Development **127**(11): 2421-2432.

Lo Celso, C., D. M. Prowse, et al. (2004). "Transient activation of beta-catenin signalling in adult mouse epidermis is sufficient to induce new hair follicles but continuous activation is required to maintain hair follicle tumours." Development **131**(8): 1787-1799.

Logan, C. Y. and R. Nusse (2004). "The Wnt signaling pathway in development and disease." Annu Rev Cell Dev Biol **20**: 781-810.

Lowry, W. E., C. Blanpain, et al. (2005). "Defining the impact of beta-catenin/Tcf transactivation on epithelial stem cells." Genes Dev **19**(13): 1596-1611.

Luetteke, N. C., H. K. Phillips, et al. (1994). "The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase." Genes Dev **8**(4): 399-413.

Luetteke, N. C., T. H. Qiu, et al. (1993). "TGF alpha deficiency results in hair follicle and eye abnormalities in targeted and waved-1 mice." Cell **73**(2): 263-278.

Mann, G. B., K. J. Fowler, et al. (1993). "Mice with a null mutation of the TGF alpha gene have abnormal skin architecture, wavy hair, and curly whiskers and often develop corneal inflammation." Cell **73**(2): 249-261.

Matise, M. P., D. J. Epstein, et al. (1998). "Gli2 is required for induction of floor plate and adjacent cells, but not most ventral neurons in the mouse central nervous system." Development **125**(15): 2759-2770.

Matise, M. P. and A. L. Joyner (1999). "Gli genes in development and cancer." Oncogene **18**(55): 7852-7859.

Merrill, B. J., U. Gat, et al. (2001). "Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin." Genes Dev **15**(13): 1688-1705.

Merrill, B. J., H. A. Pasolli, et al. (2004). "Tcf3: a transcriptional regulator of axis induction in the early embryo." Development **131**(2): 263-274.

Mikkola, M. L., J. Pispa, et al. (1999). "Ectodysplasin, a protein required for epithelial morphogenesis, is a novel TNF homologue and promotes cell-matrix adhesion." Mech Dev **88**(2): 133-146.

Mill, P., R. Mo, et al. (2003). "Sonic hedgehog-dependent activation of Gli2 is essential for embryonic hair follicle development." Genes Dev **17**(2): 282-294.

Millar, S. E., K. Willert, et al. (1999). "WNT signaling in the control of hair growth and structure." Dev Biol **207**(1): 133-149.

Mills, A. A., B. Zheng, et al. (1999). "p63 is a p53 homologue required for limb and epidermal morphogenesis." Nature **398**(6729): 708-713.

Ming Kwan, K., A. G. Li, et al. (2004). "Essential roles of BMPR-IA signaling in differentiation and growth of hair follicles and in skin tumorigenesis." Genesis **39**(1): 10-25.

Muller-Rover, S., B. Handjiski, et al. (2001). "A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages." J Invest Dermatol **117**(1): 3-15.

Munz, B., H. Smola, et al. (1999). "Overexpression of activin A in the skin of transgenic mice reveals new activities of activin in epidermal morphogenesis, dermal fibrosis and wound repair." EMBO J **18**(19): 5205-5215.

Murillas, R., F. Larcher, et al. (1995). "Expression of a dominant negative mutant of epidermal growth factor receptor in the epidermis of transgenic mice elicits striking alterations in hair follicle development and skin structure." EMBO J **14**(21): 5216-5223.

Nakamura, M., M. M. Matzuk, et al. (2003). "Control of pelage hair follicle development and cycling by complex interactions between follistatin and activin." FASEB J **17**(3): 497-499.

Nakamura, M., J. P. Sundberg, et al. (2001). "Mutant laboratory mice with abnormalities in hair follicle morphogenesis, cycling, and/or structure: annotated tables." Exp Dermatol **10**(6): 369-390.

Nguyen, H., B. J. Merrill, et al. (2009). "Tcf3 and Tcf4 are essential for long-term homeostasis of skin epithelia." Nat Genet **41**(10): 1068-1075.

Nguyen, H., M. Rendl, et al. (2006). "Tcf3 governs stem cell features and represses cell fate determination in skin." Cell **127**(1): 171-183.

Nijhof, J. G., K. M. Braun, et al. (2006). "The cell-surface marker MTS24 identifies a novel population of follicular keratinocytes with characteristics of progenitor cells." Development **133**(15): 3027-3037.

Nilsson, M., A. B. Unden, et al. (2000). "Induction of basal cell carcinomas and trichoepitheliomas in mice overexpressing GLI-1." Proc Natl Acad Sci U S A **97**(7): 3438-3443.

Nowak, J. A., L. Polak, et al. (2008). "Hair follicle stem cells are specified and function in early skin morphogenesis." Cell Stem Cell **3**(1): 33-43.

Oro, A. E., K. M. Higgins, et al. (1997). "Basal cell carcinomas in mice overexpressing sonic hedgehog." Science **276**(5313): 817-821.

Oshimori, N. and E. Fuchs (2012). "Paracrine TGF-beta signaling counterbalances BMP-mediated repression in hair follicle stem cell activation." Cell Stem Cell **10**(1): 63-75.

Osorio, K. M., S. E. Lee, et al. (2008). "Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation." Development **135**(6): 1059-1068.

Osorio, K. M., K. C. Lilja, et al. (2011). "Runx1 modulates adult hair follicle stem cell emergence and maintenance from distinct embryonic skin compartments." J Cell Biol **193**(1): 235-250.

Pan, Y., M. H. Lin, et al. (2004). "gamma-secretase functions through Notch signaling to maintain skin appendages but is not required for their patterning or initial morphogenesis." Dev Cell **7**(5): 731-743.

Park, H. L., C. Bai, et al. (2000). "Mouse Gli1 mutants are viable but have defects in SHH signaling in combination with a Gli2 mutation." Development **127**(8): 1593-1605.

Pawson, T. and J. Schlessingert (1993). "SH2 and SH3 domains." Curr Biol **3**(7): 434-442.

Pelton, R. W., B. Saxena, et al. (1991). "Immunohistochemical localization of TGF beta 1, TGF beta 2, and TGF beta 3 in the mouse embryo: expression patterns

suggest multiple roles during embryonic development." J Cell Biol **115**(4): 1091-1105.

Plikus, M. V. (2012). "New activators and inhibitors in the hair cycle clock: targeting stem cells' state of competence." J Invest Dermatol **132**(5): 1321-1324.

Plikus, M. V., R. E. Baker, et al. (2011). "Self-organizing and stochastic behaviors during the regeneration of hair stem cells." Science **332**(6029): 586-589.

Plikus, M. V., J. A. Mayer, et al. (2008). "Cyclic dermal BMP signalling regulates stem cell activation during hair regeneration." Nature **451**(7176): 340-344.

Rabbani, P., M. Takeo, et al. (2011). "Coordinated activation of Wnt in epithelial and melanocyte stem cells initiates pigmented hair regeneration." Cell **145**(6): 941-955.

Reddy, S., T. Andl, et al. (2001). "Characterization of Wnt gene expression in developing and postnatal hair follicles and identification of Wnt5a as a target of Sonic hedgehog in hair follicle morphogenesis." Mech Dev **107**(1-2): 69-82.

Rendl, M., L. Lewis, et al. (2005). "Molecular dissection of mesenchymal-epithelial interactions in the hair follicle." PLoS Biol **3**(11): e331.

Rhee, H., L. Polak, et al. (2006). "Lhx2 maintains stem cell character in hair follicles." Science **312**(5782): 1946-1949.

Schmidt-Ullrich, R., T. Aebischer, et al. (2001). "Requirement of NF-kappaB/Rel for the development of hair follicles and other epidermal appendices." Development **128**(19): 3843-3853.

Schmidt-Ullrich, R. and R. Paus (2005). "Molecular principles of hair follicle induction and morphogenesis." Bioessays **27**(3): 247-261.

Shi, Y. and J. Massague (2003). "Mechanisms of TGF-beta signaling from cell membrane to the nucleus." Cell **113**(6): 685-700.

Silva-Vargas, V., C. Lo Celso, et al. (2005). "Beta-catenin and Hedgehog signal strength can specify number and location of hair follicles in adult epidermis without recruitment of bulge stem cells." Dev Cell **9**(1): 121-131.

Snippert, H. J., A. Haegebarth, et al. (2010). "Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin." Science **327**(5971): 1385-1389.

St-Jacques, B., H. R. Dassule, et al. (1998). "Sonic hedgehog signaling is essential for hair development." Curr Biol **8**(19): 1058-1068.

Takahashi, K. and S. Yamanaka (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." Cell **126**(4): 663-676.

Tumbar, T. (2012). "Ontogeny and Homeostasis of Adult Epithelial Skin Stem Cells." Stem Cell Reviews and Reports **8**(2): 561-576.

Ullrich, A. and J. Schlessinger (1990). "Signal transduction by receptors with tyrosine kinase activity." Cell **61**(2): 203-212.

Uyttendaele, H., A. A. Panteleyev, et al. (2004). "Activation of Notch1 in the hair follicle leads to cell-fate switch and Mohawk alopecia." Differentiation **72**(8): 396-409.

Van Mater, D., F. T. Kolligs, et al. (2003). "Transient activation of beta -catenin signaling in cutaneous keratinocytes is sufficient to trigger the active growth phase of the hair cycle in mice." Genes Dev **17**(10): 1219-1224.

Vangenderen, C., R. M. Okamura, et al. (1994). "Development of Several Organs That Require Inductive Epithelial-Mesenchymal Interactions Is Impaired in Lef-1-Deficient Mice." Genes Dev **8**(22): 2691-2703.

Vidal, V. P., M. C. Chaboissier, et al. (2005). "Sox9 is essential for outer root sheath differentiation and the formation of the hair stem cell compartment." Curr Biol **15**(15): 1340-1351.

Waghmare, S. K., R. Bansal, et al. (2008). "Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells." EMBO J **27**(9): 1309-1320.

Wang, L. C., Z. Y. Liu, et al. (2000). "Regular articles: conditional disruption of hedgehog signaling pathway defines its critical role in hair development and regeneration." J Invest Dermatol **114**(5): 901-908.

Watt, F. M., S. Estrach, et al. (2008). "Epidermal Notch signalling: differentiation, cancer and adhesion." Curr Opin Cell Biol **20**(2): 171-179.

Werner, S., H. Smola, et al. (1994). "The function of KGF in morphogenesis of epithelium and reepithelialization of wounds." Science **266**(5186): 819-822.

Willert, K. and R. Nusse (1998). "Beta-catenin: a key mediator of Wnt signaling." Curr Opin Genet Dev **8**(1): 95-102.

Xie, J., M. Murone, et al. (1998). "Activating Smoothened mutations in sporadic basal-cell carcinoma." Nature **391**(6662): 90-92.

Yamasaki, M., A. Miyake, et al. (1996). "Structure and expression of the rat mRNA encoding a novel member of the fibroblast growth factor family." J Biol Chem **271**(27): 15918-15921.

Yang, A., R. Schweitzer, et al. (1999). "p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development." Nature **398**(6729): 714-718.

Yano, K., L. F. Brown, et al. (2001). "Control of hair growth and follicle size by VEGF-mediated angiogenesis." J Clin Invest **107**(4): 409-417.

Yuhki, M., M. Yamada, et al. (2004). "BMPR1A signaling is necessary for hair follicle cycling and hair shaft differentiation in mice." Development **131**(8): 1825-1833.

Zhang, Y., T. Andl, et al. (2008). "Activation of beta-catenin signaling programs embryonic epidermis to hair follicle fate." Development **135**(12): 2161-2172.

Zhang, Y. V., J. Cheong, et al. (2009). "Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells." Cell Stem Cell **5**(3): 267-278.

Zhang, Y. V., B. S. White, et al. (2010). "Stem cell dynamics in mouse hair follicles: a story from cell division counting and single cell lineage tracing." Cell Cycle **9**(8): 1504-1510.

Zhou, P. B., C. Byrne, et al. (1995). "Lymphoid Enhancer Factor-1 Directs Hair Follicle Patterning and Epithelial-Cell Fate." Genes Dev **9**(6): 700-713.

CHAPTER 3

RUNX1 AND P21 SYNERGISTICALLY LIMIT THE EXTENT OF HAIR FOLLICLE STEM CELL QUIESCENCE *IN VIVO*[2]

3.1 Summary

Mechanisms of tissue stem cell (SC) quiescence control are important for normal homeostasis and for preventing cancer. Cyclin-dependent kinase inhibitors (CDKis) are known inhibitors of cell cycle progression. We document CDKis expression *in vivo* during hair follicle stem cell (HFSC) homeostasis and find p21 (cyclin-dependent kinase inhibitor 1a, Cdkn1a), p57, and p15 up-regulated at quiescence onset. p21 appears important for HFSC timely onset of quiescence. Conversely, we find that Runx1 (runt related transcription factor 1), which is known for promoting HFSC proliferation, represses p21, p27, p57, and p15 transcription in HFSC *in vivo*. Intriguingly, in cell culture, tumors, and normal homeostasis, Runx1 and p21 interplay modulates proliferation in opposing directions under the different conditions. Unexpectedly, Runx1 and p21 synergistically limit the extent of HFSC quiescence *in vivo*, which antagonizes the role of p21 as a cell cycle inhibitor. Importantly, we find in cultured keratinocytes that Runx1 and p21 bind to the p15 promoter and synergistically repress p15 mRNA transcription, thereby restraining cell

[2] This chapter has been published in Proc Natl Acad Sci USA and is reprinted with permission. (Lee J, Hoi CS, Lilja KC, White BS, Lee SE, Shalloway D, Tumbar T. (2013). Proc Natl Acad Sci U S A. 110(12):4634-4639). Author contributions: S.E.L provided cell line reagents for Figures 3.2B-C and sorted cell samples for some QRT-PCR experiments; B.S.W. and D.S performed mathematical and computational modeling of experimental data shown in Figures 3.3C-D, F-J, Figure 3.4, Figure 3.5A-C, and Figure 3.11; C.S.L.H and K.C.L performed experiments and analyzed data for Figures 3.7 and 3.8A-B; J.L performed and analyzed all other experiments; J.L. and T.T. designed research and wrote the paper.

cycle arrest. This documents a surprising ability of a CDKi (p21) to act as a direct transcriptional repressor of another CDKi (p15). We unveil a robust *in vivo* mechanism that enforces quiescence of HFSCs, and a context-dependent role of a CDKi (p21) to limit quiescence of SCs, potentially by directly down-regulating mRNA levels of (an)other CDKi(s).

3.2 Materials and Methods

Mice

All mice were treated according to Cornell University Institutional Animal Care and Use Committee protocols. To generate H2B-GFP and p21 (*Cdkn1a*) mice, we crossed each of *K5-tTA* (CD1) and *pTRE-H2BGFP* (Tumbar, Guasch et al. 2004) (FVB) mice to *Cdkn1a−/−* mice (Jackson Laboratories strain B6;129S2-Cdkn1a$^{tm1Tyj/J}$, #003263). The *K14-Cre* (CD1, from E. Fuchs, Rockefeller University, New York, NY) line was mated with *Runx1*$^{fl/fl}$ (C57BL/6, from N. Speck, University of Pennsylvania, Philadelphia, PA), which was further mated with *Cdkn1a−/−* mice to generate dKO mice and all of the different genotypes. The crossing steps were designed to keep the ratio of genetic backgrounds constant for reducing variability of the TPA/DMBA treatment. BrdU was injected intraperitoneally in PBS at 25μg/g of body weight 2hrs before mice were killed.

DMBA/TPA Treatments

DMBA (7,12-Dimethylbenz(a)anthracene)/TPA (12-O-tetradecanoylphorbol-13-acetate) treatments were carried out according to (Hennings, Glick et al. 1993). Briefly, (i) *K14-Cre+; Runx1$^{fl/+}$; Cdkn1a$^{+/+}$*, (ii), *K14-Cre+; Runx1$^{fl/+}$; Cdkn1a$^{-/-}$*, (iii) *K14-Cre+; Runx1$^{fl/fl}$; Cdkn1a$^{+/+}$*, and (iv) *K14-Cre+; Runx1$^{fl/fl}$; Cdkn1a$^{-/-}$* mice were obtained by parallel crosses to keep a constant genetic background mix. Their back skins were treated with a single dose of 32μg of DMBA in 200μL of acetone at second telogen [between postnatal day (PD) 60 and PD 90]. Then we biweekly applied 12.4μg/body weight of TPA dissolved in 200μL of acetone per each treatment for 20wk (total 40 TPA treatments). The hair was shaved for efficient delivery of TPA as well as to count the number of papillomas. Mice were monitored weekly for papilloma counts.

H2B-GFP Pulse-Chase System

H2B-GFP fusion protein is expressed under Keratin 5 (K5)-tTA control, which is shut off by feeding the mice with doxycycline (chase). Upon cell division, H2B-GFP intensity is halved, allowing division counting during the chase. H2BGFP mRNA production is tightly controlled upon doxycycline administration. See also (Tumbar, Guasch et al. 2004; Waghmare, Bansal et al. 2008; Zhang, Cheong et al. 2009) for more details.

Fluorescence Activated Cell Sorting

To isolate different cell populations (bulge and non-bulge), mouse back and belly skin cells were isolated using trypsin digestion and filtered through 70-μm and

40-μm cell strainers (BD Biosciences) to obtain single-cell suspensions. Cells were labeled with CD34-Biotin (eBiosciences), Streptavidin-APC (BD Pharmingen), and anti–α6-integrin-PE (CD49f; BD Pharmingen). Propidium iodide (PI; Sigma) was used to label the dead cells. Bulge cells (CD34+/α6-integrin+) and nonbulge cells (CD34−/α6-integrin+) were isolated (Tani, Morris et al. 2000; Trempus, Morris et al. 2003) for culture, RNA extraction, and H2B-GFP profile acquisition. BD fluorescence-activated cell sorting (FACS) Aria was used for cell sorting and BD LSRII was used for cell cycle analysis. Both machines were used to obtain H2B-GFP profiles from doxycycline (doxy)-chased mice at different stages.

RNA Extraction, Quantitative Real-Time PCR, and Primer Sequences

The RNeasy Micro kit (Qiagen) was used to extract RNA from FACS-isolated cells. All RNA samples were quality checked with Bioanalyzer (Bio-Rad) through the Cornell Life Sciences Core Laboratories Center. cDNA synthesis was carried out using the iScript (Bio-Rad) cDNA synthesis kit according to the manufacturer's instruction. Quantitative real-time (q) PCR was carried out using homebrew SYBR Green PCR buffer. The iCycler (Bio-Rad) PCR machine was used to detect fluorescence. Primers used for qPCR are listed in Table 3.1.

Chromatin Immunoprecipitation

For Runx1 chromatin immunoprecipitation (ChIP) on p21 promoter, ~0.5–1 × 107 of wild-type (*Runx1*$^{fl/fl}$) and inducible Runx1 knockout (iKO) (*K14-CreERT2* x *Runx1*$^{fl/fl}$) keratinocytes were treated with 4-OHT (4-Hydroxytamoxifen; Sigma-

Aldrich) for 4 d before fixing. For Runx1 and p21 ChIP on p15 promoter, primary keratinocytes between passages three and seven, originally isolated from newborn skin, were used. All steps were carried out according to (Ortt and Sinha 2010) with minor variations. Cells were fixed for 3 min. Sonic Dismembrator model 100 (Fisher Scientific) was used to sonicate the chromatin. Protein A agarose beads (Millipore, #16-125) were used for immunoprecipitation. H3 (Abcam, #1791), Runx1 (Abcam, ab23980), p21 (Abcam, ab2961), H3K27me3 (Millipore, #07-449), and IgG (Imaginex or Cell Signaling) antibodies were used for immunoprecipitation. The PCR purification kit (Qiagen) was used to purify ChIP DNA. ChIP signals were calculated using ChIP-qPCR. Briefly, ChIP signals were interpolated from serial dilution of input DNA standards. Then, fold enrichment of (i) Runx1 signal over IgG (for Runx1 ChIP on p21 promoter) or (ii) WT Runx1 or p21 signals over respective double-(d)KO signals was compared at different regions of p21 or p15 promoters. Primers for ChIP-qPCR are listed in Table 3.1.

Cell Lines and Transfections

For ChIP of H3K27me3 (Figure 3.2C), mouse keratinocytes between passages three and seven, originally isolated from newborn skin, were used (Osorio, Lilja et al. 2011). Cells were either transfected with K14-Runx1-myc (epitope) or K14-eGFP-myc (epitope) constructs using previously described method (Hoi, Lee et al. 2010), or overexpressed with inducible Runx1 transgene for 3hrs. Cells were fixed and harvested for ChIP analysis.

For CMV-p15-GFP transfection (Figure 3.9 D and E and Figure 3.8D), primary keratinocytes were isolated from newborn pups of WT, p21 KO, and Runx1/p21 dKO mice. At ~70% confluency, 100 μL of serum-free medium containing a mixture of either 1 μg of CMVeGFP- myc (Osorio, Lilja et al. 2011) or CMV-p15-GFP (Origene) with 6 μL of Trans-Fectin lipid reagent (Bio-Rad) was added to the dishes. Cells were pulsed with BrdU and harvested between 24 and 30 h of incubation for cell cycle analysis, using the APC-BrdU Flow Kit (BD Biosciences) according to the manufacturer's protocol. Cells were analyzed by BD-LSRII flow cytometry and gated specifically for transfected (GFP+) cells.

For knockdown, 250 μL of serum-free media containing 12 μL of TransFectin lipid reagent and 4 μg of either p15-shRNA-GFP (Origene) or noneffective scrambled-shRNA-GFP (Origene) were added to 50~70% confluent WT, p21 KO, and Runx1/p21 dKO primary keratinocytes. Cells were harvested ~72 h posttransfection and the cell cycle was profiled using the same approach as above.

Immunofluorescence Staining, Paraffin Staining, Microscopy, and Image Processing

Optimal Cutting Temperature compound-embedded frozen skin blocks were cryosectioned, fixed, immunoblocked, and incubated with antibodies. Antibodies and dilutions were BrdU (1:300; Abcam), Caspase-3 (1:500; R&D Systems), and CD34 (1:150; BD Pharmingen). For paraffin staining, paraffin sections were deparaffinized in toluol, rehydrated with gradients of ethanol: 100%, 90%, and 70%. Antigen

retrieval was carried out in either EDTA (Invitrogen) or citrate buffer (Invitrogen) by microwaving 10 min twice. Sections were blocked and incubated with antibodies. Antibodies and dilutions used were as follows: p21, 1:50 (Abcam); and K14, 1:2,000 (Abcam). A fluorescence light microscope (Nikon) equipped with a charge-coupled device 12-bit digital camera (Retiga EXi; QImaging) was used to analyze all stained slides.

Western Blot

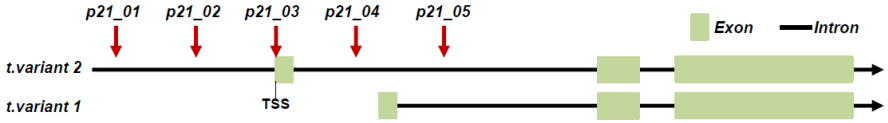
WT, p21 KO, Runx1 KO, and Runx1/p21 dKO mice were killed at PD21. Protein extracts were made via a previously described method (Osorio, Lee et al. 2008). Anti-p15 (Abcam; 1:500) and anti-Actin (Millipore; 1:20,000) were used for immunoblotting.

Computation and Mathematical Modeling

We infer the proportions of cells that have divided n times by a hierarchical Bayesian analysis that incorporates both the biological variations between mice, the experimental variations that broaden the FACS histograms into Gaussians, and the variations between the extents of H2B-GFP labeling in different mice. In addition, an enhanced method of variational Bayesian Gaussian mixture modeling (Attias 1999; Bishop 2006) is developed that can cope with noise in the low fluorescence region and identify peaks that are only partially resolved. This method may find applicability in FACS data analysis (Boedigheimer and Ferbas 2008) (Pyne, Hu et al. 2009) and a wide range of other onedimensional Gaussian mixture modeling problems as it

provides a Bayesian method for computing a combined posterior estimate from multiple mixture modeling analyses that systematically includes both intra- and intermixture variations (See detailed description in Figure 3.11).

Table 3.1. Primers for ChIP-qPCR / qPCR

Application	Gene/Region	Forward	Reverse
ChIP-qPCR	*p21_01*	TCACACCACAACAGACAAGGA	TCAGGTCGTCAGGTTAGATGG
ChIP-qPCR	*p21_02*	CTCCTCCCACAAGGTCAAAC	TCTTACCCACTGAGCCATCTC
ChIP-qPCR	*p21_03*	AGACCAGCAGCAAAATCGGA	TTGGGTATCATCAGGTCTCCACC
ChIP-qPCR	*p21_04*	GCGGCTGTTTTTCTTGGTAG	AGACGAGGAAAGCAGTTCCA
ChIP-qPCR	*p21_05*	GGGATTCCAGGTGCGGCTCG	TCGGAGAACCGCTCTGCGCT
ChIP-qPCR	*Gapdh*	ACCGAAGAACAACGAGGAGA	GAGAACAGGGGAAATGGAGA
		Cdkn1a gene / promoter	
			
ChIP-qPCR	*p15_01*	ACTTCAGGATTACAACCAAAGC	GGCGGTATTTCACAGTTAGCA
ChIP-qPCR	*p15_02*	GAAGGGGAGAAGCAAAGGAC	CCAAATGAGGGAAAGAAATACAA
ChIP-qPCR	*p15_03*	GACTTTCTGTGCCTCTGTTCTCTA	TGGGCTCAAGTTTGGAATATG
ChIP-qPCR	*Wnt4_01 (neg)* *(Devgan, Mammucari et al. 2005)*	GTCTCAGCAGCAGGCTTACC	TCAAATCCTGGATGCCAATC
ChIP-qPCR	*Wnt4_02 (pos)* *(Devgan, Mammucari et al. 2005)*	CAATTCCCAGCACCAAAAGT	TCCAAATAAGGTAGGTACCCAAAG
qPCR	*p15*	TCCACAGGCTAAATGGGAAA	TCAGAATCCAGGCATCAAGG
qPCR	*p16*	GCCGCACCGGAATCCT	TGCACCGTAGTTGAGCAGAAGA
qPCR	*p18*	AAAGATGGAACTGGTTTTGCTGT	TGTGCTTCATAAGGAACTCCAC
qPCR	*p19*	GCCACTGTCTCCAGCCTTAC	CCCCAAACACACACACTCAA
qPCR	*p21*	GTCCAATCCTGGTGATGTCC	CAGACGAAGTTGCCCTCCA
qPCR	*p27*	TCTCTTCGGCCCGGTCAAT	GGGGCTTATGATTCTGAAAGTCG
qPCR	*p57*	AGGGGAATGGTTGTTGAGTAA	GGCAGTTGGTATGGGCAGT
qPCR	*Bcl-2*	TGGGATGCCTTTGTGGAACT	ACAGCCAGGAGAAATCAAACAG
qPCR	*Bcl-xl*	GGCTGGGACACTTTTGTGGAT	GCGCTCCTGGCCTTTCC
qPCR	*Mcl1*	CCCTCCCCCATCCTAATCAG	AGTAACAATGGAAAGCATGCCAAT
qPCR	*Bim*	TTGGAGCTCTGCGGTCCTT	CAGCGGAGGTGGTGTGAAT
qPCR	*Puma*	GCGGCGGAGACAAGAAGA	AGTCCCATGAAGAGATTGTACATGAC
qPCR	*Bak*	AATGGCATCTGGACAAGGAC	GTTCCTGCTGGTGGAGGTAA
qPCR	*Bax*	TGGAGCTGCAGAGGATGATTG	AGCTGCCACCCGGAAGA

3.3 Introduction

Tight regulation of quiescence likely maintains tissue stem cells (SCs) such as blood and hair follicle (HF) and may prevent cancer (Orford and Scadden 2008; Fuchs 2009). Cyclin-dependent kinase inhibitors (CDKis) regulate the cell cycle in different tissue cell types. Some of the Cip/Kip (CDK interacting protein/kinase inhibitor protein) family [p21 (cyclin-dependent kinase inhibitor 1a, Cdkn1a), p27, and p57] and the Ink4 family (p15, p16, p18, and p19) of CDKis are known to prevent hematopoietic SC (HSC) exhaustion by limiting cell divisions (Matsumoto, Takeishi et al. 2011; Tesio and Trumpp 2011; Zou, Yoshihara et al. 2011), but their relevance in regulating other mammalian SCs *in vivo* is not well established. Some of the CDKis, such as p21, play different roles in HSCs *in vivo* vs. in stress conditions such as cell culture, transplantation, and carcinogenesis (Cheng, Rodrigues et al. 2000; van Os, Kamminga et al. 2007; Foudi, Hochedlinger et al. 2009; Viale, De Franco et al. 2009; Matsumoto, Takeishi et al. 2011; Zou, Yoshihara et al. 2011).

Interestingly, it has been recently suggested that cells may possess a sensing mechanism for the cellular level of total CDKis and that some CDKis act redundantly (Matsumoto, Takeishi et al. 2011; Tesio and Trumpp 2011; Zou, Yoshihara et al. 2011). It is well known that in general CDKis are transcriptionally regulated (Orford and Scadden 2008; Tesio and Trumpp 2011) but the SC-specific transcriptional control and how it modulates adult SC behavior *in vivo* are not well understood. Because some SCs are apparently cycling frequently, it is also unclear how robust quiescence control actually is *in vivo*.

Here we use the mouse HF as a classical quiescent SC mammalian system (Cotsarelis 2006; Waghmare, Bansal et al. 2008) to examine transcriptional regulation and the role of CDKis, in an attempt to decipher the robustness of hair follicle stem cell (HFSC) cell cycle control in normal and stress conditions. HF adult homeostasis (hair cycle) is composed of distinct and synchronous phases of SC self-renewal, differentiation, and quiescence (Figure 3.1A) (Tumbar 2012). In each hair cycle upon activation from the quiescent (telogen) to the growth phase (anagen), a subset of bulge SCs migrate out into the hair germ below without self-renewal to differentiate and produce hair shafts (Zhang, Cheong et al. 2009). The remaining bulge cells self-renew at anagen and reenter quiescence at catagen when the matrix and the differentiated cells undergo apoptosis and the HFs return to quiescence (Tumbar 2012).

Transcription factors such as NFATc1 (nuclear factor of activated T cells, cytoplasmic, calcineurin-dependent 1), Sox9 (SRY-box containing gene 9), Lhx2 (LIM homeobox protein 2), and Runx1 (runt related transcription factor 1) (Tumbar 2012) regulate the proliferation and maintenance of HFSCs during adult homeostasis. In particular Runx1, a transcription factor of the Runt family member of cancer genes (Speck and Gilliland 2002), promotes HFSC activation from quiescence to migrate, proliferate, and differentiate at telogen (Osorio, Lee et al. 2008) and self-renew at anagen (Hoi, Lee et al. 2010) and controls timely emergence of HFSCs and proper maturation during embryogenesis (Osorio, Lilja et al. 2011). Moreover, Runx1 is indispensable for skin tumorigenesis and for keratinocyte growth in culture. Interestingly, Runx1 deletion in mouse skin epithelium up-regulates p21 in the HFSCs (Hoi, Lee et al. 2010). p21 is a tumor supressor downstream of p53 known to arrest the

cell cycle in response to DNA damage and telomere shortening (Choudhury, Ju et al. 2007). p21 knockout (KO) results in increased skin squamous papilloma in response to TPA (12-O-tetradecanoylphorbol-13-acetate)/DMBA (7,12-Dimethyl-benz(a)-anthracene) (Topley, Okuyama et al. 1999), although it did not affect the progression to malignancy (Weinberg, Fernandez-Salas et al. 1999). Moreover, p21 KO mouse keratinocytes showed increased proliferation *in vitro* and enhanced short-term engraftment ability *in vivo* (Topley, Okuyama et al. 1999). Runx1 interacts genetically with p21 in cultured keratinocytes (Hoi, Lee et al. 2010), but the significance of this interaction in the normal tissue or in skin tumorigenesis is unknown.

Here we document the transcriptional regulation of CDKis and their potential control by Runx1 in HFSCs.We use mouse genetics tools to concomitantly target Runx1 and p21 *in vivo* and examine their interplay in tumors, in cells in culture, and in the normal hair cycle.

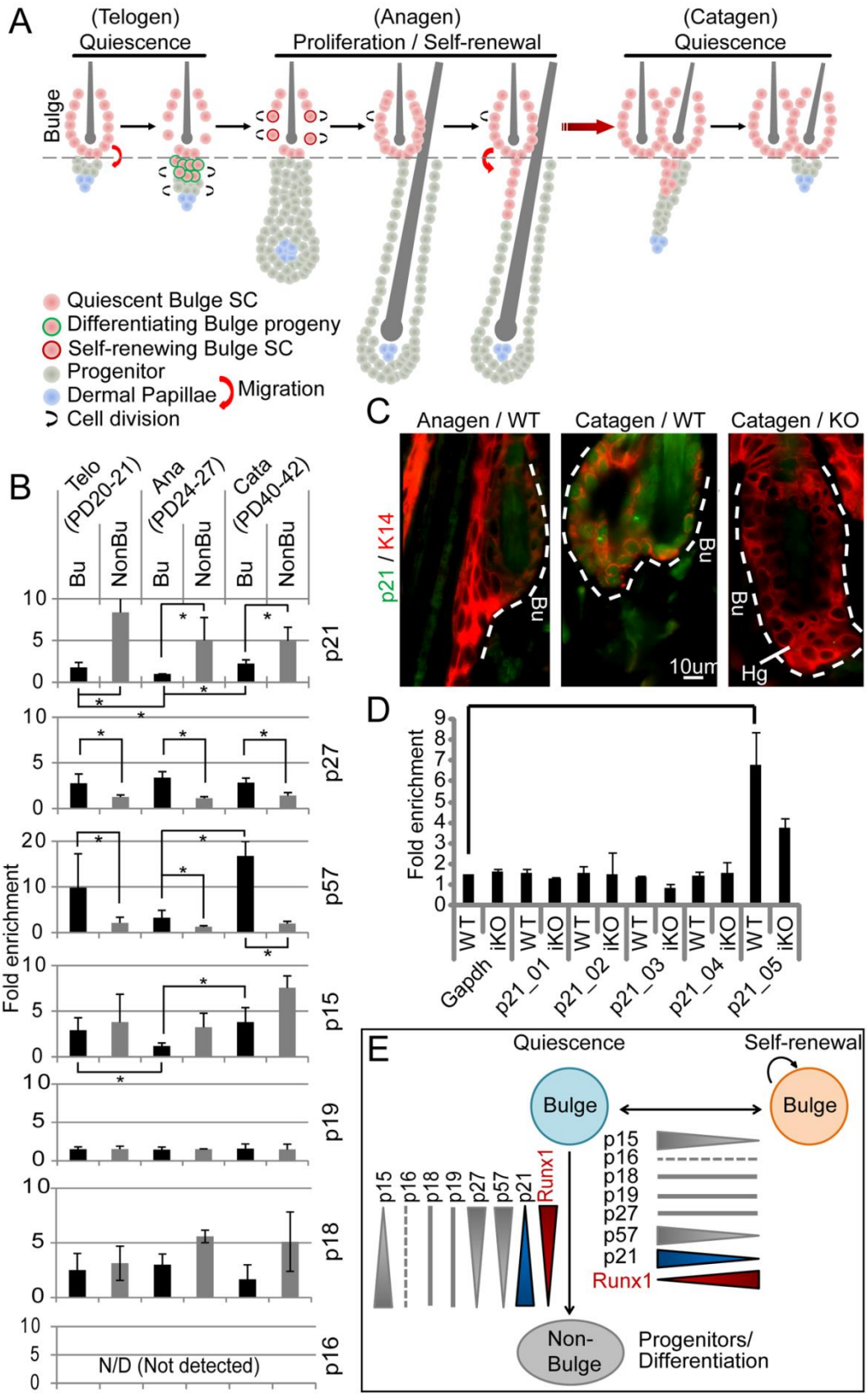
3.4 Results

3.4.1 Regulation of CDKi expression during normal HFSC homeostasis

To explore CDKis transcriptional control in HFSCs, we documented their mRNA expression profiles in CD34+/α6-integrin+ bulge cells isolated from mouse skin by fluorescence-activated cell sorting (FACS) at different hair cycle stages. We used HFs at proliferation (anagen) vs. quiescence (catagen/telogen) and performed quantitative real-time (q)PCR. First, p21 mRNA expression was up-regulated more than twofold in the WT quiescent vs. proliferative bulge (Figure 3.1B), in agreement

Figure 3.1. Regulation of p21 and other CDKis expression in normal HFSC homeostasis.

(A) HFSC dynamics during normal skin homeostasis. (B) qPCR analysis of CDKi mRNAs in WT [telogen (PD20–21), anagen (PD24–27), and catagen (PD40–42)] bulge (CD34+/α6+, Bu) and nonbulge (CD34−/α6+, NonBu) cells (Figure 3.2A) (C) Immuno-fluorescence staining of paraffin skin sections at anagen (PD27) and catagen (PD39) shows nuclear p21 (green) in quiescence (KO: p21KO). (D) ChIP-qPCR with Runx1 antibody and primers to Gapdh and different regions of the p21 promoter (Table 3.1, page 88) in WT and Runx1 inducible KO (iKO) keratinocytes (Figure 3.2 B and C). (E) Model for cell cycle control factors show inverse relation of p21 and other factors with Runx1 protein expression (red) (Osorio, Lee et al. 2008). (Two relative comparisons are independent).

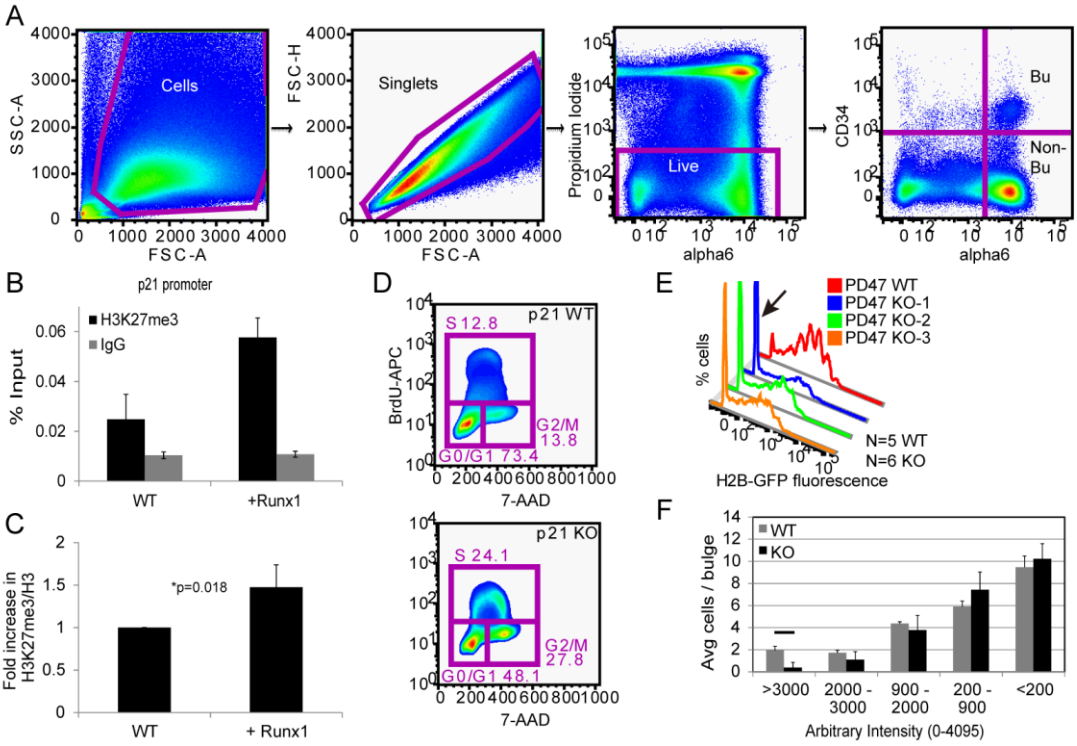

A
(Telogen)
Quiescence
(Anagen)
Proliferation / Self-renewal
(Catagen)
Quiescence
Bulge
Quiescent Bulge SC
Differentiating Bulge progeny
Self-renewing Bulge SC
Progenitor
Dermal Papillae
Migration
Cell division
B
Telo (PD20-21)
Ana (PD24-27)
Cata (PD40-42)
Bu
NonBu
Fold enrichment
p21
p27
p57
p15
p19
p18
p16
N/D (Not detected)
C
Anagen / WT
Catagen / WT
Catagen / KO
p21 / K14
Bu
Hg
10um
D
Fold enrichment
WT
iKO
Gapdh
p21_01
p21_02
p21_03
p21_04
p21_05
E
Quiescence
Self-renewal
Bulge
Bulge
p15
p16
p18
p19
p27
p57
p21
Runx1
Non-Bulge
Progenitors/ Differentiation

with nuclear p21 protein signal in catagen/telogen bulge (Figure 3.1C). Nuclear p21 is known to inhibit cell proliferation (Zhou, Liao et al. 2001). Intriguingly, p57 and p27 involved in HSC quiescence (Matsumoto, Takeishi et al. 2011; Zou, Yoshihara et al. 2011), were up-regulated in the bulge relative to nonbulge (CD34−/α6+) at all hair cycle stages (Figure 3.1B), suggesting that they may be responsible for the overall low rates of bulge cell division (Waghmare, Bansal et al. 2008). Moreover, p57 and p15 mRNAs were found at higher levels in the catagen and telogen bulge than in the anagen. p19 was equally expressed and p18 seemed to be more specific to the nonbulge cells, whereas p16 was undetectable in any of the cell fractions tested (Figure 3.1B).

Next we wondered what regulates transcription of these CDKis throughout the hair cycle. We previously showed that Runx1 protein is expressed in the bulge during anagen and is at low levels at catagen and telogen when HFSCs are quiescent (Osorio, Lee et al. 2008). This suggested an inverse correlation of Runx1 protein and CDKi mRNA level in the bulge throughout the hair cycle, making Runx1 a potential repressor of CDKis (Figure 3.1E). Indeed, p21 was up-regulated in the bulge when we previously deleted Runx1 at anagen (Hoi, Lee et al. 2010). Here we isolated CD34+/α6+ bulge cells from Runx1 KO mice at quiescence [postnatal day (PD) 21] before the onset of the Runx1 phenotype. We observed not only p21, but also p27 (three of three mice) and p57 (two of three mice showed increased expression whereas one showed a similar level compared with WT) mRNAs up-regulated in the Runx1 KO (CD34+/α6+) cells relative to WT (see Figure 3.9A). This was consistent with the

Figure 3.2. Runx1 directly represses p21 promoter and the p21 knockout bulge cells proliferate more *in vitro* and *in vivo*.

(A) FACS plots of bulge and nonbulge skin cell isolation. Single cells freshly isolated from mouse skin were stained and immediately FACS sorted. Gating order is cells → singlets → live cells (PI negative) → bulge cells, α6-integrin+ (PE)/CD34+ (APC); nonbulge cells, PE+/APC−. (B and C) ChIP-qPCR with H3K27me3, H3, and IgG antibodies on p21 promoter upon Runx1 overexpression (Materials and Methods). Note the increased H3K27me3 on p21 promoter upon Runx1 overexpression. (D) Representative cell cycle plot of BrdU-APC and 7-AAD in p21 WT and KO. (E) Representative H2B-GFP FACS histograms of PD21–47 doxy-chased bulge cells. (F) Microscopy quantification of H2B-GFP fluorescence in bulge cells.


A
SSC-A
FSC-A
Cells
FSC-H
Singlets
Propidium Iodide
Live
CD34
Bu
Non-Bu
alpha6
B
p21 promoter
H3K27me3
IgG
% Input
WT
+Runx1
C
Fold increase in H3K27me3/H3
*p=0.018
WT
+ Runx1
D
p21 WT
S 12.8
G2/M 13.8
G0/G1 73.4
BrdU-APC
7-AAD
p21 KO
S 24.1
G2/M 27.8
G0/G1 48.1
E
PD47 WT
PD47 KO-1
PD47 KO-2
PD47 KO-3
% cells
N=5 WT
N=6 KO
H2B-GFP fluorescence
F
WT
KO
Avg cells / bulge
>3000
2000 - 3000
900 - 2000
200 - 900
<200
Arbitrary Intensity (0-4095)

prolonged quiescence of HFSCs in the Runx1 KO and suggested that Runx1 acts as a repressor for several CDKis *in vivo*.

Loss of Runx1 in cell culture induced a severe proliferation defect and p21 loss rescued this phenotype (Hoi, Lee et al. 2010), suggesting an important genetic interaction between Runx1 and p21 *in vitro*. To ask whether this interaction works by the direct binding of Runx1 to the p21 promoter, which Runx1 may inhibit, we performed chromatin immunoprecipitation (ChIP) on mouse keratinocytes. This revealed enrichment of Runx1 at a conserved Runx1 binding site in the p21 genetic locus (p21_05) relative to other regions [p21_01-04 and Gapdh (glyceraldehyde-3-phosphate dehydrogenase)] and relative to Runx1 inducible KO cells (Figure 3.1D). In addition, the p21 promoter was enriched with a repressive histone mark, H3K27me3 upon Runx1 overexpression (Figure 3.2 B and C), likely via recruitment of corepressors of the Polycomb group proteins (Yu, Mazor et al. 2012). In conclusion, a number of CDKis, including p21, p27, p57, and p15, are expressed and/or up-regulated at the mRNA level in the bulge at quiescence. Their redundant presence likely contributes to overall bulge cell quiescence and timely cell cycle exit in catagen. Moreover, p21 is likely a direct downstream repressional target of Runx1 and downregulation of Runx1 protein in the bulge could result in the up-regulation of p21, followed by cell cycle exit at catagen.

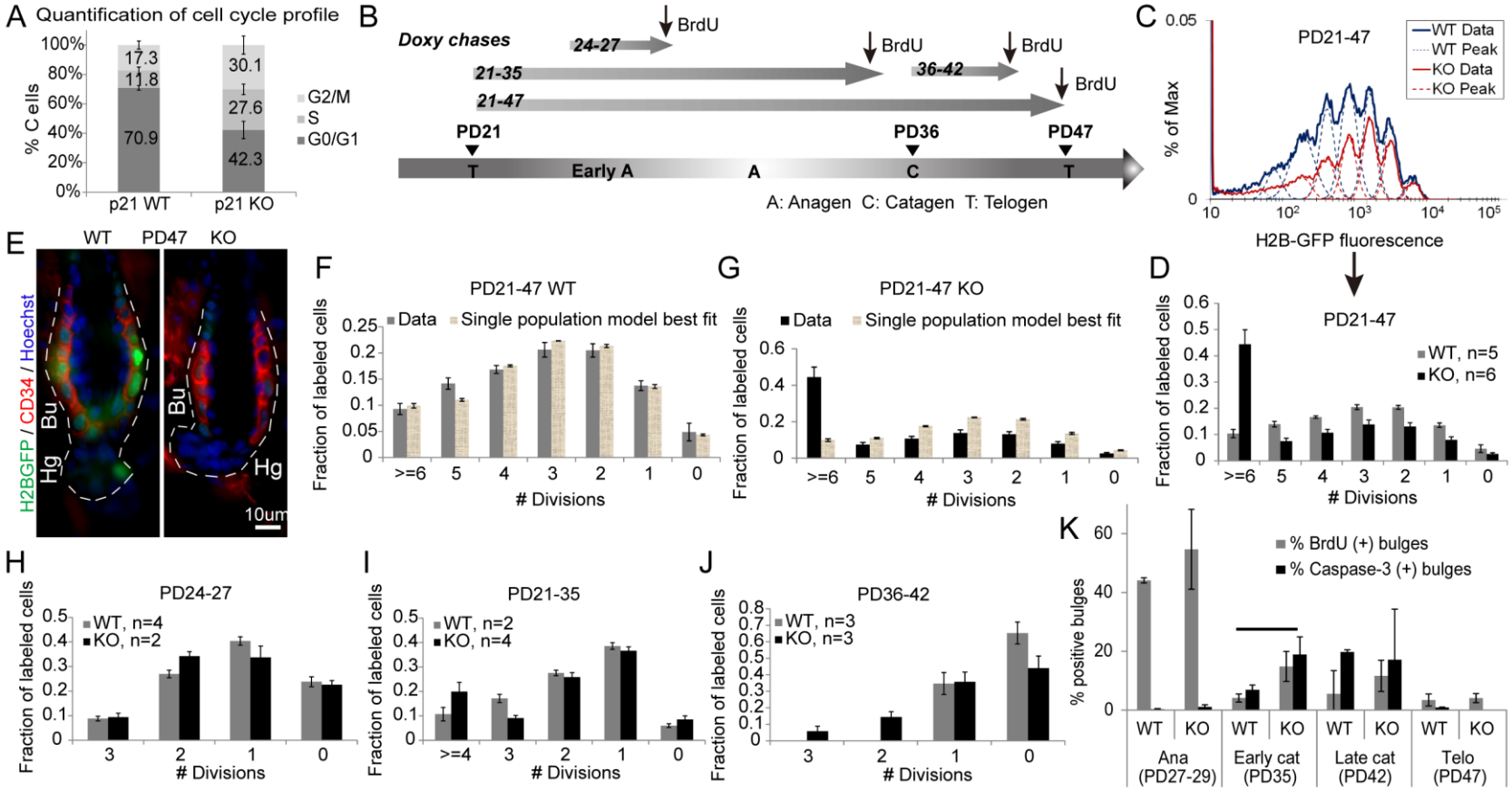
3.4.2 P21 reduces proliferation of HFSCs in vitro and controls their timely exit into quiescence in vivo

The role of p21 in HFSCs *in vivo* has not been yet analyzed. p21 KO mice are viable and fertile and have roughly normal appearance. To understand the role of p21 in HFSCs, we analyzed the proliferation ability of bulge (CD34+/α6+) HFSCs from WT and p21 KO cultured cells (Figure 3.3A and Figure 3.2D). This revealed increased S and G2/M phase with a concomitant decrease of G0/G1 p21 KO cells, suggesting that p21 indeed plays a role in bulge cell proliferation *in vitro*, as we expected from its role in keratinocytes (Topley, Okuyama et al. 1999).

To understand whether p21 loss affects bulge cell proliferation and skin homeostasis *in vivo* we used our previously developed H2B-GFP tet-repressible pulse-chase system (Tumbar, Guasch et al. 2004; Waghmare, Bansal et al. 2008). Doxycycline (doxy) shuts down H2B-GFP expression and results in dilution of the label at each division by twofold. We performed doxy "chases" on mice from PD21 to PD47 to span the first adult hair cycle (Figure 3.3B), killed the mice, and examined CD34+/α6+ bulge cell divisions by FACS (Figure 3.3C and Figure 3.2E). A computational method was devised to delineate individual peaks of H2B-GFP fluorescence (FL) from FACS (Figure 3.3C and Figure 3.4A). This quantification revealed that p21 KO bulge cells divided more than WT during one normal hair cycle, as seen in the increased cell fraction with lower H2B-GFP FL (Figure 3.3 C and D). In addition, the bright H2B-GFP label-retaining cells (LRCs), which were previously deemed a population of relatively unused "reserve" SCs (Hsu, Pasolli et al. 2011), were also induced to replicate by the p21 KO (Figure 3.3E and Figure 3.2F). However,

Figure 3.3. p21 represses bulge cell proliferation *in vitro* and *in vivo* and regulates timely cell cycle exit.

(A) CD34+/α6+ sorted cells cultured and BrdU pulsed for 30 min were labeled with anti–BrdU-APC and 7-AAD, and the cell cycle was profiled using a FACS analyzer (at least two mice per condition, n = 3 experiments). (B) Doxy chase schemes followed by BrdU injection 2 h before mice were killed. (C) Individual H2B-GFP peak modeling of PD21–47 bulge cell FACS data (Figures 3.2E and 3.4A and Materials and Methods). (D) Quantification of the fraction of labeled bulge cells under each peak from C. (E) Representative HF image of PD21–47 doxy-chased bulge cells stained with CD34. (F and G) PD21–47 WT (F) and KO (G) H2B-GFP FACS data fitted with a homogeneous population Poisson model (Figure 3.4B and Materials and Methods). Note KO data are not fitted. Error bars show the 1σ credible intervals (data) or the predicted proportions corresponding to the 1σ intervals of the best-fit replication rates (model). (H–J) Fraction of labeled bulge cells from different doxy "chased" mice. (K) Quantification of percentage of BrdU+ bulges (gray) and percentage of caspase-3+ bulges (black) in skin sections from mice at age or stage indicated. Two to three mice per genotype and 30–300 HFs per mouse were analyzed.


A Quantification of cell cycle profile
% C ells
100%
80%
60%
40%
20%
0%
17.3
11.8
70.9
30.1
27.6
42.3
p21 WT
p21 KO
G2/M
S
G0/G1
B
Doxy chases
24-27
21-35
21-47
BrdU
36-42
PD21
PD36
PD47
T
Early A
A
C
T
A: Anagen C: Catagen T: Telogen
C
PD21-47
WT Data
WT Peak
KO Data
KO Peak
% of Max
0.05
0
10
10²
10³
10⁴
10⁵
H2B-GFP fluorescence
D
PD21-47
WT, n=5
KO, n=6
Fraction of labeled cells
Divisions
E
WT
PD47
KO
H2BGFP / CD34 / Hoechst
Bu
Hg
10um
F
PD21-47 WT
Data
Single population model best fit
Fraction of labeled cells
Divisions
G
PD21-47 KO
Data
Single population model best fit
Fraction of labeled cells
Divisions
H
PD24-27
WT, n=4
KO, n=2
Fraction of labeled cells
Divisions
I
PD21-35
WT, n=2
KO, n=4
Fraction of labeled cells
Divisions
J
PD36-42
WT, n=3
KO, n=3
Fraction of labeled cells
Divisions
K
% BrdU (+) bulges
% Caspase-3 (+) bulges
% positive bulges
WT
KO
Ana (PD27-29)
Early cat (PD35)
Late cat (PD42)
Telo (PD47)

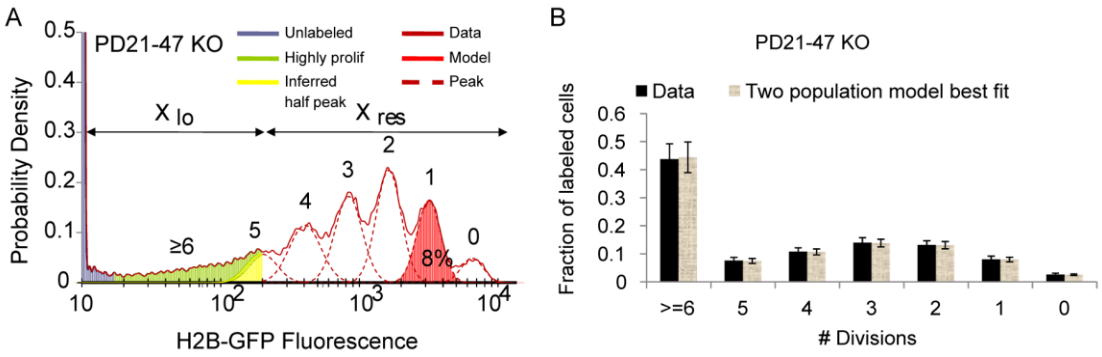
by PD47, WT and p21 KO mice showed a similar fraction (3–4%) of BrdU+ bulges, suggesting that the majority of HFs of both types reentered quiescence (Figure 3.3K, PD47). Interestingly, all CDKi levels remained unchanged in p21 KO, suggesting that enough CDKis may be already expressed in the bulge to maintain bulge cell quiescence even in the absence of p21 (Figure 3.5F).

A simple Poisson model fitted the WT experimental data (Figure 3.3F) (Materials and Methods) and indicated approximately three divisions per bulge cell in one hair cycle, as we previously reported (Waghmare, Bansal et al. 2008). Intriguingly, this model did not fit the KO experimental data indicating that the bulge cells do not have a homogeneous response to p21 KO (Figure 3.3G and Materials and Methods). A mathematical model assuming that only a subpopulation of the bulge cells respond to p21 loss does fit the data (Figure 3.4B) and suggests that ~40% of the KO bulge cells present at PD47 had replicated at least three times more than WT. Collectively, p21 KO bulge cells divide more *in vitro* and *in vivo*. Extra divisions affect even the most quiescent bulge cells, but all cells eventually return to quiescence.

To see whether p21 regulates rates of proliferation or timely cell cycle exit during quiescence, we performed short doxy chases at anagen and catagen (Figure 3.3B). Bulge cells self-renew symmetrically at anagen to replenish their pool depleted by migration in the preceding phase (Zhang, Cheong et al. 2009; Zhang, White et al. 2010). Analysis of sorted bulge cells showed comparable numbers of bulge cell divisions in WT and p21 KO skin during PD24–27 (rapid self-renewal phase) (Figure 3.3H and Figure 3.5A) with no significant differences in BrdU+ bulges at PD27 (Figure 3.3K), suggesting that p21 did not affect the rates of self-renewal in early

Figure 3.4. Statistical and theoretical analysis.

(A) Quantification of the proportion of cells within each H2B-GFP peak by variational Bayes Gaussian mixture modeling. (B) Best fit model of the PD21–47 KO data using the two-population Poisson model. The error bars show the standard deviations of the predicted proportions. See Materials and Methods for details.


A
PD21-47 KO
Unlabeled
Highly prolif
Inferred half peak
Data
Model
Peak
X lo
X res
Probability Density
H2B-GFP Fluorescence
≥6
5
4
3
2
1
0
8%
B
PD21-47 KO
Data
Two population model best fit
Fraction of labeled cells
Divisions
>=6
5
4
3
2
1
0

anagen. However, by PD35, which marks the beginning of catagen when bulge cells return to quiescence, increased growth of ~9% (as estimated by mathematical modeling) of p21 KO bulge cells could be detected from FACS and BrdU data at PD35 (Figure 3.3 I and K and Figure 3.5 B and D). This increase in p21 KO bulge cell proliferation was more pronounced by PD42 as seen in additional divisions clearly detected in the PD36–42 chase (Figure 3.3J and Figure 3.5C). Our mathematical modeling indicated that the WT bulge cells reduced their division rates by sevenfold in catagen compared with anagen (Figure 3.11), but that the reduction for KO cells was roughly half as much. Both WT and KO bulge cells eventually entered quiescence by telogen as shown by BrdU staining (Figure 3.3K, PD47). In summary, whereas WT bulge cells divided approximately three times in anagen and had greatly reduced proliferation after catagen onset, p21 KO bulge cells overall divided approximately one to two more times before entering quiescence, reflecting the existence of a cell fraction with a higher replication rate and/or extended replication period.

Interestingly, the extra bulge cells generated by prolonged proliferation due to p21 loss did not increase the bulge cell pool (Figure 3.3E and Figure 3.6A) measured by number of CD34+ cells per bulge at second telogen. This is, at least in part, explained by a KO-correlated increase in apoptosis in early catagen. Levels of apoptotic caspase-3+ cells in the p21 KO bulge showed an approximately threefold increase at PD35, the stage when we also detected failure of p21 KO cells to enter quiescence (Figure 3.3K and Figure 3.5E). This was consistent with the known anti-

Figure 3.5. p21 loss delays the onset of bulge cell quiescence in normal homeostasis.

(A–C) Individual H2B-GFP peak modeling of PD24–27, PD21–35, and PD36–42 doxy-chased bulge cells FACS data. Note in C that p21 KO bulge cells fail to enter quiescence in a timely manner and undergo additional divisions. (D and E) Representative hair follicle (HF) images of immunofluorescent-labeled skin sections for BrdU (D) and Caspase-3 (E). Arrowhead indicates a positive cell. (F) Ratio of fold mRNA qPCR enrichment of p21 KO over WT at PD42 shows insignificant changes. At least two mice per genotype were used.

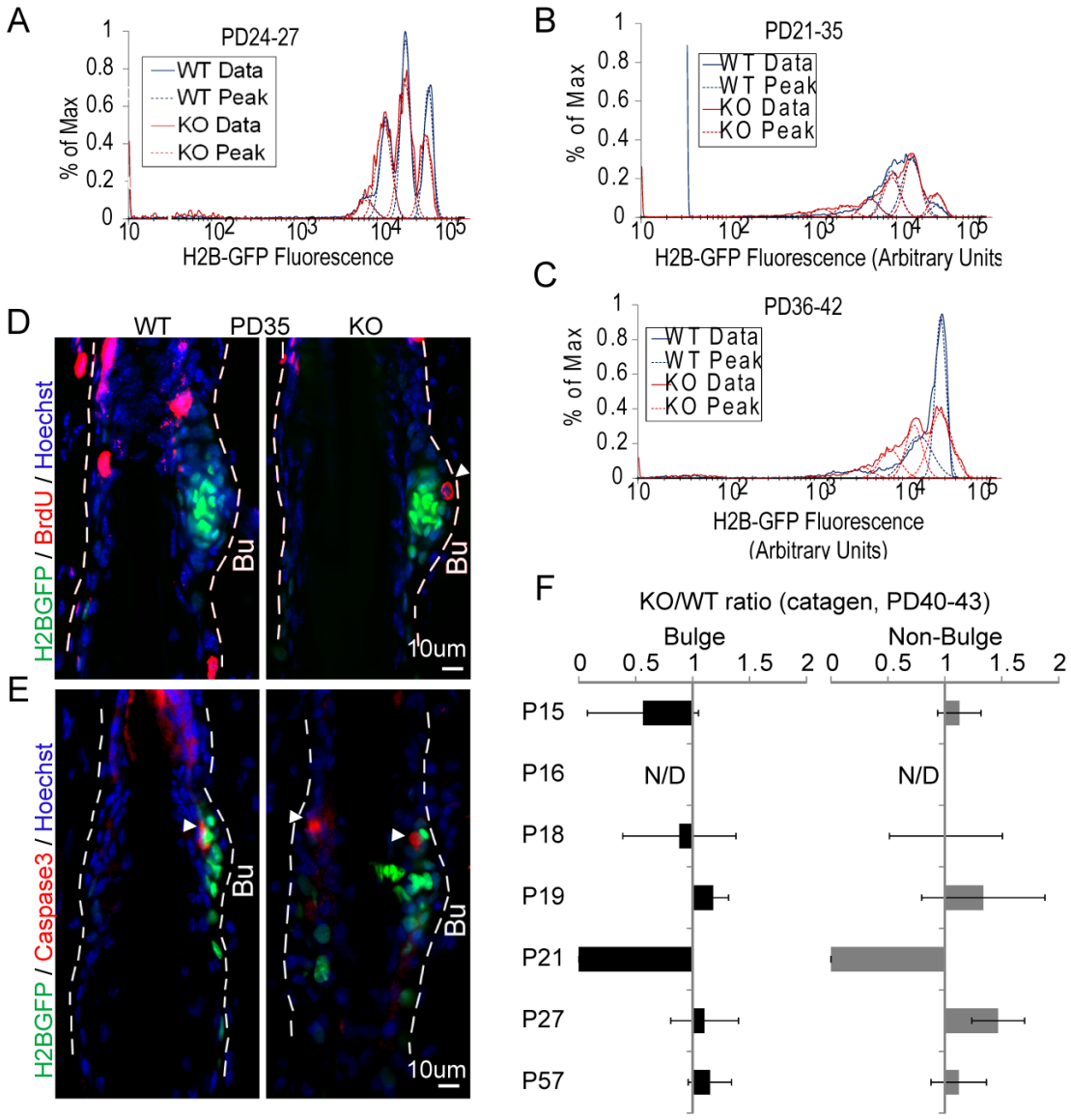
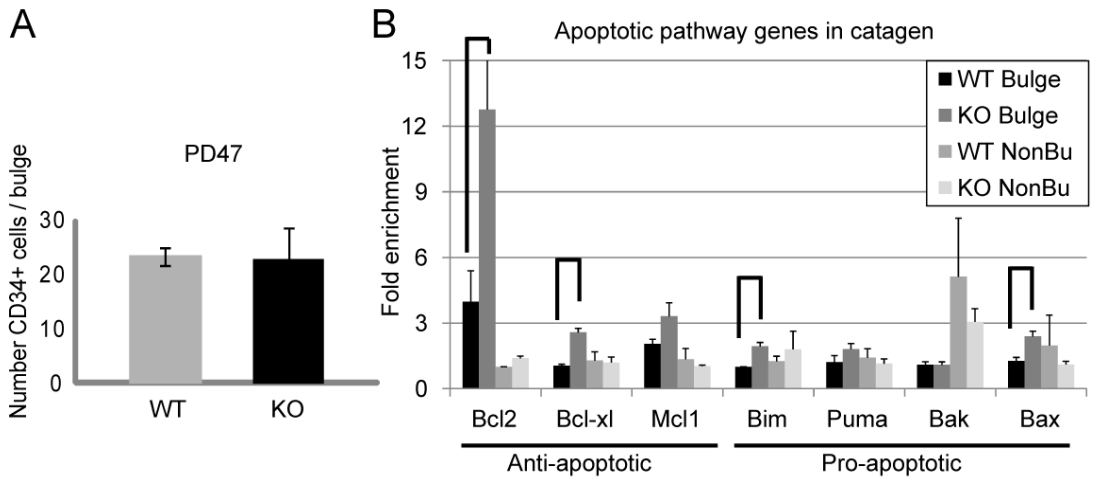

A
PD24-27
WT Data
WT Peak
KO Data
KO Peak
% of Max
H2B-GFP Fluorescence
B
PD21-35
WT Data
WT Peak
KO Data
KO Peak
% of Max
H2B-GFP Fluorescence (Arbitrary Units
C
PD36-42
WT Data
WT Peak
KO Data
KO Peak
% of Max
H2B-GFP Fluorescence
(Arbitrary Units)
D
WT
PD35
KO
H2BGFP / BrdU / Hoechst
Bu
10um
E
H2BGFP / Caspase3 / Hoechst
Bu
10um
F
KO/WT ratio (catagen, PD40-43)
Bulge
Non-Bulge
P15
P16
P18
P19
P21
P27
P57
N/D
N/D

Figure 3.6. p21 knockout bulge cell pool size is maintained constant by increased apoptosis.

(A) Quantification in immunofluorescently labeled skin sections of CD34+ cells per bulge in p21 WT and KO at second telogen. Note the bulge cell pool size did not increase in the KO. (B) qPCR analysis of apoptotic genes at PD42. At least two different mice were used per genotype. Note the significant differences in expression between WT and p21 KO (black lines).


A
PD47
Number CD34+ cells / bulge
30
20
10
0
WT
KO
B
Apoptotic pathway genes in catagen
Fold enrichment
15
12
9
6
3
0
WT Bulge
KO Bulge
WT NonBu
KO NonBu
Bcl2
Bcl-xl
Mcl1
Bim
Puma
Bak
Bax
Anti-apoptotic
Pro-apoptotic

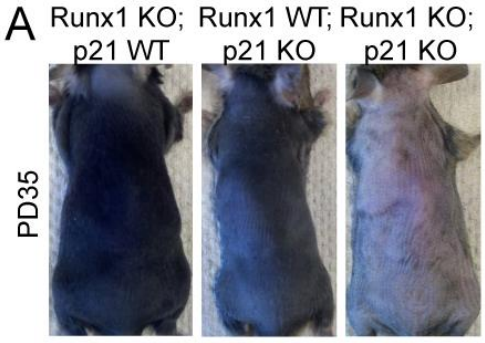
apoptotic role of p21 (Suzuki, Tsutomi et al. 1998). Moreover, catagen bulge up-regulated not only pro- but also anti-apoptotic genes (Willis, Day et al. 2003), likely to protect against excessive bulge cell death upon p21 loss (Figure 3.6B). Finally, we did not observe colocalization of BrdU+ and caspase-3+ bulge cells, ruling out the possibility that p21 KO bulge cells die as they enter the cell cycle at an unpermitted time. Thus, it appears that some p21 KO bulge cells die to compensate for prolonged proliferation and to keep the bulge cell pool size constant. In conclusion, we found that p21 is important for regulating general rates of HFSCs *in vitro* whereas *in vivo* its up-regulation at quiescence, likely driven by down-regulation of Runx1 protein in the bulge, limits the extent of HFSC divisions during homeostasis.

3.4.3 Runx1/p21 genetic interaction is context dependent

Although p21 KO rescued the proliferation phenotype of keratinoyctes in culture (Hoi, Lee et al. 2010), it was unclear whether Runx1-driven effects on the cell cycle *in vivo* or in tumors were mediated by p21. To ask whether prolonged HFSC quiescence in Runx1 single-KO mice is mediated via de-repression of p21, we generated Runx1/p21 double-KO (dKO) mice and examined the proliferation onset by skin color change (pink to black). Unexpectedly, HFSCs in the dKO mice not only failed to overcome the prolonged quiescence imposed by Runx1 loss as we predicted, but in fact remained in telogen much longer (Figure 3.7A–C, Figure 3.10A). These *in vivo* data contrasted strikingly with our previous *in vitro* study (Hoi, Lee et al. 2010) and indicated that *in vivo* but not *in vitro*, double KO of a CDKi (p21) and its upstream repressor (Runx1) triggered deep repression of the bulge cell proliferation.

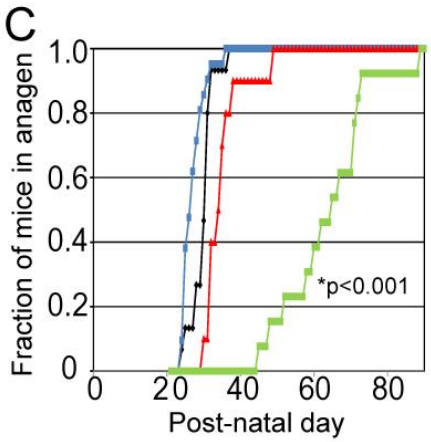
Figure 3.7. Runx1–p21 interaction plays context-dependent roles in regulating proliferation.

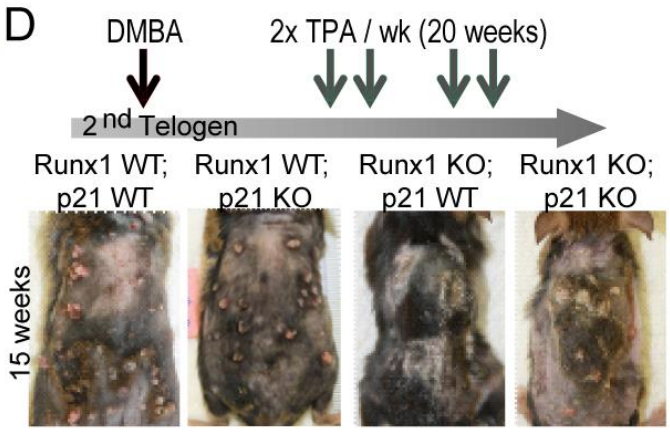
(A) Skin coat color at PD35 indicates different hair cycle stages: Single p21 and Runx1 KO mice were in anagen (black); dKO was in telogen (pink). (B) Summary of mice and color coding used for data in C–F. (C) Anagen onset is delayed in Runx1 KO and much further delayed in dKO mice ($p<0.001$, Log-Rank and Wilcoxon tests) (D) (Upper) Two-step carcinogenesis scheme (Materials and Methods); (Lower) Representative mouse pictures at 15 wk TPA treatment. Treatment began upon second telogen onset in all mice. (E) Number of papillomas over time with t-test statistical analysis. Note: additional statistical test using standard least square regression (shown in Figure 3.10) revealed no statistical differences between Runx1 KO and dKO during the entire 20 weeks ($p=0.3506$) (dKO vs p21 KO, $p=0.0004$; dKO vs WT, $p=0.0006$). However, between weeks 13-16, Runx1 KO and dKO mice showed a significant difference ($p=0.0229$) (F) Fraction of papilloma-bearing mice (statistical analyses for E and F also in Figure 3.8 A and B). (dKO vs WT $p=0.00628$; dKO vs p21KO $p=0.0325$; dKO vs Runx1KO $p=0.0631$).

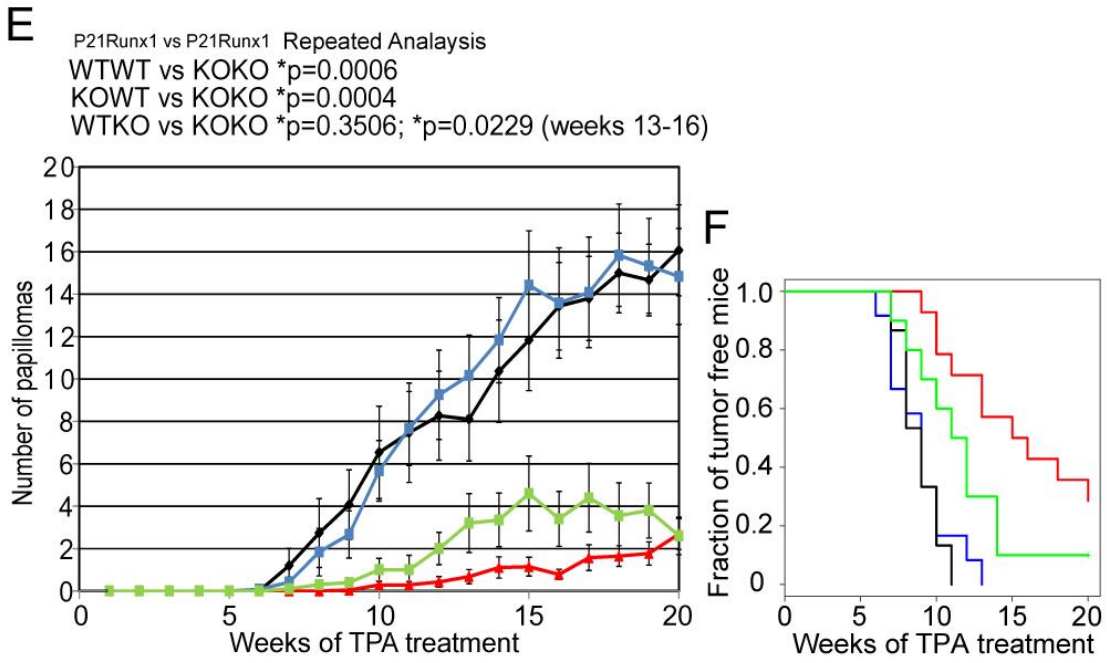


B

Genotype			# mice	
Color	Runx1	p21	Fig 3C	Fig 3E&3F
	+/+	+/+	15	15
	+/+	-/-	21	12
	Δ/Δ	+/+	10	14
	Δ/Δ	-/-	13	10







To examine the effect of Runx1 and p21 interaction in tumors we used a two-step carcinogenesis protocol, using DMBA/TPA on single- and double-Runx1/p21 KO mice (Figure 3.7D). This generates skin tumors that eventually progress to squamous cell carcinoma (SCC) (Kemp 2005). At least some of these SCCs were previously shown to originate from bulge SCs (Lapouge, Youssef et al. 2011; White, Tran et al. 2011; Scheitz, Lee et al. 2012). Because carcinogenic treatment can produce different amounts of tumors on different genetic backgrounds (Hennings, Glick et al. 1993), we controlled the mating scheme to generate four different genotypes with a fixed ratio of BL6 and CD1 backgrounds for the final experimental mice. This may be responsible for the high error bars that did not allow us to detect significant tumor differences between WT and p21 KO mice, as previously reported (Topley, Okuyama et al. 1999; Weinberg, Fernandez-Salas et al. 1999). Upon DMBA/TPA treatment initiated at the second telogen, Runx1 KO mice, however, had significantly ($P < 0.05$, weeks 11–20) fewer papillomas compared with WT, as expected. At 15–16 wk of TPA treatment, the dKO mice had significantly more tumors than the Runx1 KO ($P < 0.05$) but fewer than WT and p21 KO mice (Figure 3.7 D–F, Figure 3.10 B), suggesting a partial rescue of the tumor impairment phenotype imposed by the Runx1 KO. However, by the 20th week of TPA treatment the differences in tumor burden or fraction of tumor-free mice were not statistically different from those of the Runx1 single-KO mice. These results suggest that p21 repression via Runx1 plays a minor role in promoting skin papillomas and that additional Runx1 targets must be at play for full tumor cell growth *in vivo*. In conclusion, we found three different responses of skin epithelial cells to concomitant

Figure 3.8. Context-dependent control of hair follicle stem cells (HFSCs) by Runx1–p21 interplay.

(A and B) Statistical analysis of Runx1/p21 mice: (A) Kaplan–Meier survival test for Figure 3.7F and (B) Student's t test for Figure 3.7E. (C) qPCR of control genes [Runx1, p21, and Keratin14 (epithelial-specific)] on cultured cells. (D) Cell cycle analysis of WT and dKO cells harvested 24 h after transfection with CMV-eGFP-myc or CMV-p15-GFP (details in Materials and Methods). A representative cell cycle histogram of dKO with or without p15-GFP transfection is shown. (E) p15 knockdown of dKO keratinocytes via transfection of p15shRNA-GFP or scrambled-shRNA-GFP (control) followed by cell cycle analysis. (Left) Selection of transfected cells (FITC+); (Right) representative dKO cell cycle histogram of p15-shRNA (red) and scrambled-shRNA (blue).

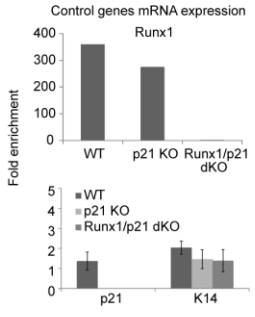
A

Latency (P-values)	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V	Chisq	degree of freedom	p-value
Genotype=Rx1KO;p21KO	10	9	5.51	2.21	3.45			
Genotype=Rx1KO;p21WT	14	10	13.49	0.903	3.45	3.5	1	0.0631
Genotype=Rx1KO;p21KO	10	9	13.1	1.29	4.57			
Genotype=Rx1WT;p21KO	12	12	7.9	2.13	4.57	4.6	1	0.0325
Genotype=Rx1KO;p21WT	14	10	16.99	2.87	16.5			
Genotype=Rx1WT;p21KO	12	12	5.01	9.74	16.5	16.5	1	4.96E-05
Genotype=Rx1KO;p21KO	10	9	14.24	1.93	7.47			
Genotype=Rx1WT;p21WT	15	15	9.76	2.81	7.47	7.5	1	0.00628
Genotype=Rx1KO;p21WT	14	10	18.25	3.73	19.9			
Genotype=Rx1WT;p21WT	15	15	6.75	10.1	19.9	19.9	1	8.18E-06
Genotype=Rx1WT;p21KO	12	12	12.9	0.0585	0.177			
Genotype=Rx1WT;p21WT	15	15	14.1	0.0533	0.177	0.2	1	0.674
Genotype=Rx1KO;p21KO	10	9	10.25	0.152	0.233			
Genotype=Rx1KO;p21WT	14	10	21.71	6.314	16.276			
Genotype=Rx1WT;p21KO	12	12	6.39	4.917	7.086			
Genotype=Rx1WT;p21WT	15	15	7.65	7.055	11.168	26.7	3	6.87E-06

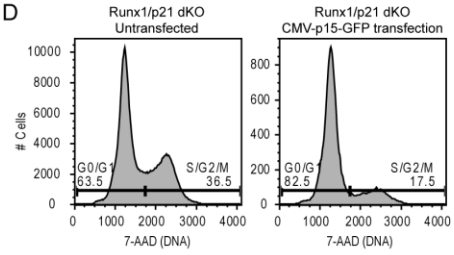
B

T-test	Weeks																			
P21/Rx1 vs P21/Rx1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
WTWT vs KOWT	0.291	0.291	0.291	0.291	0.291	0.22	0.792	0.843	0.673	0.849	0.883	0.697	0.423	0.621	0.466	0.919	0.875	0.741	0.733	0.793
WTWT vs WTKO	N/A	N/A	N/A	N/A	N/A	N/A	0.179	0.128	0.029	0.012	0.008	0.001	0.001	0.001	3E-04	5E-06	7E-06	6E-07	5E-07	1E-05
WTWT vs KOKO	N/A	N/A	N/A	N/A	N/A	N/A	0.3	0.253	0.093	0.061	0.044	0.032	0.088	0.041	0.041	0.001	0.003	3E-04	1E-04	8E-05
KOWT vs WTKO	N/A	N/A	N/A	N/A	N/A	0.289	0.095	0.113	0.025	9E-04	3E-04	3E-04	5E-05	3E-05	3E-05	3E-05	6E-05	4E-06	6E-06	4E-05
KOWT vs KOKO	0.907	0.907	0.907	0.907	0.907	0.992	0.818	0.386	0.204	0.031	0.011	0.016	0.02	0.007	0.012	0.007	0.012	0.002	0.002	1E-03
WTKO vs KOKO	N/A	N/A	N/A	N/A	N/A	N/A	0.245	0.107	0.068	0.159	0.249	0.036	0.053	0.079	0.038	0.029	0.08	0.195	0.134	0.937

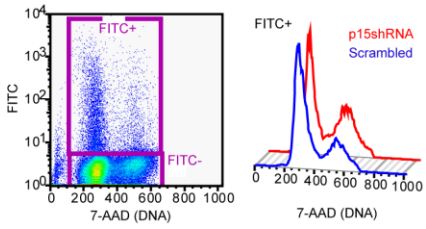
C



D



E



loss of Runx1 and p21. When in cell culture keratinocytes proliferate nearly as WT cells (Hoi, Lee et al. 2010), when in tumors they proliferate more than Runx1 single-KO cells but not as well as the WT cells, whereas when in their intact niche (bulge) HFSCs became even more profoundly quiescent upon concomitant Runx1/p21 loss.

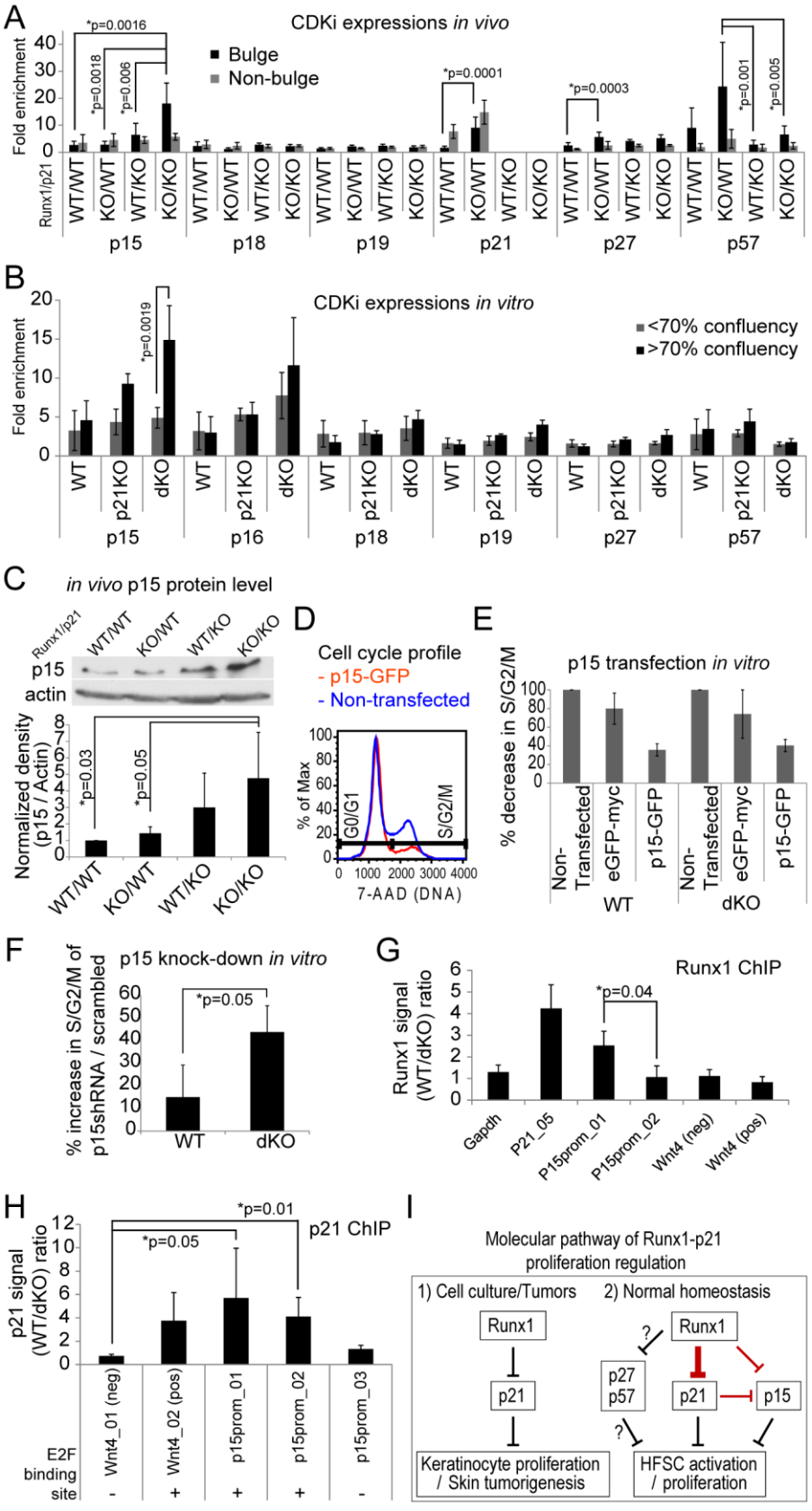
3.4.4 Runx1-p21 interplay synergistically up-regulates p15

Next, we examined the discrepancy in Runx1–p21 interplay on epithelial cells found in normal tissue conditions or in cell culture. We hypothesized that the prolonged HFSC quiescence *in vivo* might be due to compensation by redundant CDKis up-regulated upon concomitant Runx1/p21 loss. Thus, we analyzed CDKi mRNA expression in bulge (CD34+/α6+) cells from telogen mice of all four genotypes (Figure 3.9A). Interestingly, dKO bulge cells showed an approximately fivefold increase in p15 expression whereas none of the single-KO or the nonbulge cells showed p15 up-regulation *in vivo*. Moreover, this increase in p15 was also detected at the protein level (Figure 3.9C). Interestingly, we noted that p27 and p57, which were somewhat up-regulated in the single Runx1 KO, were expressed at normal levels in dKO bulge cells (Figure 3.9A) (Discussion).

Because, unlike dKO bulge cells *in vivo*, which remain profoundly quiescent, the dKO cells grow normally in culture (Hoi, Lee et al. 2010), we hypothesized that p15 up-regulation might not occur *in vitro*. Indeed, dKO cells grown at log phase (<70% confluency) did not significantly up-regulate any of the CDKis mRNA compared with p21 KO and WT cells (Figure 3.9B, gray, Figure 3.8C for control genes). However, we observed that cells grown at over 70% confluency, conditions that may better

Figure 3.9. Runx1–p21 interaction involves synergistic p15 repression.

(A) Bulge (black) and nonbulge (gray) qPCR signal of all CDKis *in vivo* in telogen (PD21) ($n \geq 3$ mice, n = 3 independent experiments). (B) qPCR analysis of CDKis in cell culture (n = 2 independent cell lines, n = 3 experiments). Note the significant p15 up-regulation in >70% confluency dKO (Figure 3.8C). (C) (Upper) Representative Western blot of p15 from total skin protein extracts at first telogen; (Lower) p15 protein level normalized to actin. n = 3 independent experiments. (D and E) Cell cycle analysis of WT and dKO cells harvested 24 h after transfection with CMV-eGFP-myc or CMVp15-GFP. A representative cell cycle histogram of GFP+ (transfected) cells is shown in D and the quantification is in E (Figure 3.8D). (F) Increase in percentage of S/G2/M-phase cells upon p15 shRNA knockdown over scrambled control is shown. WT and dKO keratinocytes were harvested 72 h after transfection and the GFP+ cells were FACS analyzed (Figure 3.8E). (G and H) Runx1 and p21 ChIPs show direct binding of Runx1 and p21 on p15 promoter in keratinocytes. Gapdh and P21_05 are the same regions shown in Figure 3.1D. Wnt4_01 (neg) and Wnt4_02 (pos) regions were previously shown to bind p21 (Devgan, Mammucari et al. 2005) (n = 3 independent experiments). (I) Molecular pathway of Runx1/p21-mediated HFSC proliferation regulation. Direct (red) and potentially direct (question mark) regulatory mechanisms revealed from this study are shown.


A
CDKi expressions in vivo
Bulge
Non-bulge
Fold enrichment
Runx1/p21
WT/WT
KO/WT
WT/KO
KO/KO
p15
p18
p19
p21
p27
p57
*p=0.0016
*p=0.0018
*p=0.006
*p=0.0001
*p=0.0003
*p=0.001
*p=0.005
B
CDKi expressions in vitro
<70% confluency
>70% confluency
Fold enrichment
*p=0.0019
WT
p21KO
dKO
p15
p16
p18
p19
p27
p57
C
in vivo p15 protein level
Runx1/p21
WT/WT
KO/WT
WT/KO
KO/KO
p15
actin
Normalized density (p15 / Actin)
*p=0.03
*p=0.05
D
Cell cycle profile
- p15-GFP
- Non-transfected
% of Max
G0/G1
S/G2/M
7-AAD (DNA)
E
p15 transfection in vitro
% decrease in S/G2/M
Non-Transfected
eGFP-myc
p15-GFP
WT
dKO
F
p15 knock-down in vitro
% increase in S/G2/M of p15shRNA / scrambled
*p=0.05
WT
dKO
G
Runx1 ChIP
Runx1 signal (WT/dKO) ratio
*p=0.04
Gapdh
P21_05
P15prom_01
P15prom_02
Wnt4 (neg)
Wnt4 (pos)
H
p21 ChIP
p21 signal (WT/dKO) ratio
*p=0.01
*p=0.05
E2F binding site
Wnt4_01 (neg)
Wnt4_02 (pos)
p15prom_01
p15prom_02
p15prom_03
I
Molecular pathway of Runx1-p21 proliferation regulation
1) Cell culture/Tumors
2) Normal homeostasis
Runx1
p21
p27
p57
p15
Keratinocyte proliferation / Skin tumorigenesis
HFSC activation / proliferation

resemble the crowded bulge cells in their niche, in fact up-regulated p15 (Figure 3.9B, black).

To test whether p15 alone could block the proliferation of cycling cells in log-phase cultures, we transfected keratinocytes with p15-GFP cDNA and examined the cell cycle profiles relative to those of nontransfected and GFP-myc control transfected cells by FACS. Gating on the GFP+ (transfected) cells showed >50% reduction in S/G2/M phases in both WT and dKO keratinocytes with a concomitant increase in G0/G1 induced by p15 overexpression (Figure 3.9D and E and Figure 3.8D). On the other hand, shRNA knockdown of p15 showed increased proliferation, which was more prominent in dKO than in WT keratinocytes (Figure 3.9F and Figure 3.8E). Finally, to explore how Runx1 and p21 may synergistically repress p15 transcription, we examined their direct binding on the p15 promoter region. p21 was previously shown to bind to E2F binding sites and repress downstream transcription on the Wnt4 (wingless-related MMTV integration site 4) promoter (Devgan, Mammucari et al. 2005). Importantly, we found that Runx1 and p21 both bind to p15 promoter regions at their own respective binding motifs (Figure 3.9 G and H), suggesting that they may directly repress p15 transcription.

3.5 Discussions

Here we probed the mechanism of HFSC quiescence in normal tissue and in stress conditions of tumors and cell culture.We found several CDKis (p21, p57, p27, and p15), known to play overlapping roles in halting the cell cycle (Matsumoto,

Takeishi et al. 2011; Tesio and Trumpp 2011) transcriptionally up-regulated either in the bulge at all stages or at the onset of HF quiescence. Using pulse-chase H2B-GFP mice (Waghmare, Bansal et al. 2008) we find that p21 represses HFSC rates of proliferation *in vitro* and limits the HFSC extent of proliferation at catagen onset *in vivo*, but does not affect HFSC self-renewal rates at anagen. Overall p21 KO bulge cells undergo approximately one to two more divisions at catagen, but increased apoptosis maintains the bulge cell pool at a constant size. Mathematical modeling suggests a temporary escape from quiescence in a fraction of the KO bulge cells (Materials and Methods). Interestingly, loss of Dacapo, the Drosophila p21 homolog, also leads to an additional round of cell division before developmental arrest (de Nooij, Letendre et al. 1996). The p21 KO bulge cells eventually enter quiescence likely via concerted action of other bulge up-regulated CDKis (p27, p57, and p15).

One of the upstream players in the mechanism of CDKi control *in vivo* appears to be the transcription factor Runx1, which may repress p21, p27, p15, and p57 mRNA production and is directly bound to the p21 promoter where it promotes H3K27me3 accumulation. We propose that p21 (and possibly p57) derepression due to down-regulation of the Runx1 protein at catagen promotes the timely onset of WT HFSC quiescence and limits the expansion of their pool (this work). Conversely, Runx1 protein expression in the bulge at anagen represses p21 expression and promotes self-renewal (Hoi, Lee et al. 2010).

The Runx1/p21 dKO mice display a surprising extension of HFSC quiescence *in vivo*, whereas keratinocytes in culture grow normally (Hoi, Lee et al. 2010). Tumors showed an intermediate effect, suggesting that additional Runx1 targets, of which

Stat3 is an important player (Scheitz, Lee et al. 2012), are at play to promote tumor growth. These data exposed a robust SC control *in vivo* that seemingly limits the normal SC pool size and enforces quiescence. A potential factor in this mechanism is another CDKi, p15, which is up-regulated specifically in the bulge only in the Runx1/p21 dKO mice. Our data suggest that when cultured cells are crowded, transcriptional up-regulation of this CDKi (p15) is driven by direct synergistic derepression of its promoter via loss of p21 and Runx1. Previously, p21 and Runx1 were shown capable of acting as transcriptional repressors (Speck and Gilliland 2002; Devgan, Mammucari et al. 2005; Ferrandiz, Caraballo et al. 2012; Marques-Torrejon, Porlan et al. 2012), and we find that they bind to the predicted DNA sites on the p15 promoter. Additionally, p15 expression is able to hamper cell cycle progression of cultured keratinocytes. These data reveal the case of one CDKi as a direct transcriptional repressor of another, to finely tune the extent of quiescence (Figure 3.9I).

Because SCs in tissues need not only to be quiescent to protect their genome, but also to activate rapidly when needed, it may be that high steady-state levels of CDKis are undesirable. An intriguing mechanism to keep CDKi levels in check may involve the dual role of some CDKis as CDK inhibitor and transcriptional repressors. Recently p27 was also shown to repress transcription, although direct targeting by p27 of another CDKi did not surface from this study (Pippa, Espinosa et al. 2011). It will be interesting to know how general is the ability of one CDKi to directly repress another, such that when one is inadvertently lost, another is up-regulated to maintain proliferation control. For example, if p15 could repress p27 and p57, this would

explain why the latter two were not expressed in the context of the Runx1/p21 dKO when the level of p15 was high. Clearly there is more to learn about CDKis beyond their traditional function in CDK inhibition. Testing triple-KO p21/Runx1/p15 mice will begin to address this model *in vivo*, although additional control layers may further enforce quiescence and possibly complicate the analysis.

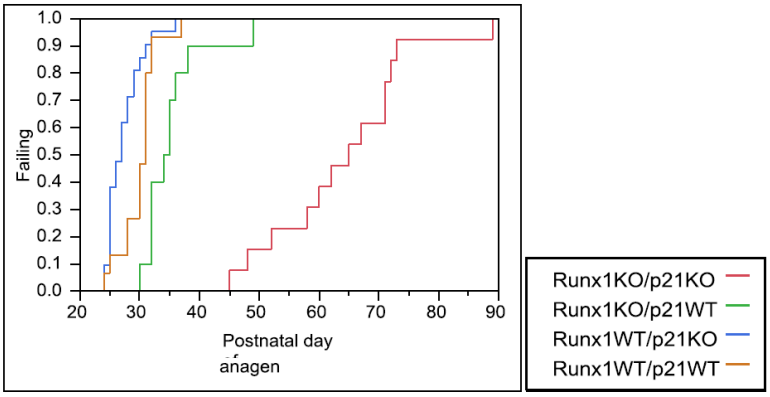
Collectively, our study uncovers a complex and robust mechanism *in vivo* that enforces SC quiescence and a constant SC pool size and is synergistically tempered by Runx1 and, unexpectedly, by its downstream CDKi target (p21). Moreover, we unveil a role of a CDKis (p21) to antagonize proliferation by direct transcriptional repression of another CDKi (p15) *in vitro*, thereby modulating the strength of cell cycle arrest.

Figure 3.10. Additional statistical analyses of Runx1/p21 interactions.

(A) Statistical analysis of anagen delay in Runx1/p21 double knockout shown in Figure 3.7C using survival test. Note both Log-Rank and Wilcoxon test show $p<0.0001$ (B) Additional statistical analysis of number of papillomas over time in Runx1/p21 mice shown in Figure 3.7E using standard least square regression approach. Comparison of Runx1 KO and dKO mice during 20 weeks duration showed no significant difference ($p=0.3506$). However, the difference was significant ($p=0.0229$) between weeks 13-16. All other comparisons showed significant differences (dKO vs p21 KO: $p=0.0004$; dKO vs WT: $p=0.0006$).

A

Product-Limit Survival Fit
Failure Plot



Time to event:
Postnatal day of anagen entrance
Grouped by
Genotype

Summary

Group	Number failed	Number censored	Mean	Std Error
Runx1KO/p21KO	13	0	64.0769	3.29619
Runx1KO/p21WT	10	0	35.3	1.69345
Runx1WT/p21KO	21	0	27.3333	0.65949
Runx1WT/p21WT	15	0	30.0667	0.78962
Combined	59	0	37.4746	2.05823

Quantiles

Group	Median Time	Lower 95%	Upper 95%	25% Failures	75% Failures
Runx1KO/p21KO	65	52	71	58	71
Runx1KO/p21WT	34.5	30	36	32	36
Runx1WT/p21KO	27	25	28	25	29
Runx1WT/p21WT	31	28	31	28	31
Combined	31	30	32	27	38

Tests Between Groups

Test	ChiSquare	DF	Prob>ChiSq
Log-Rank	60.8333	3	<.0001*
Wilcoxon	48.4068	3	<.0001*

Runx1KO/p21KO

Postnatal day of anagen entrance	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.0000	1.0000	0.0000	0.0000	0	0	13
45.0000	0.9231	0.0769	0.0739	1	0	13
48.0000	0.8462	0.1538	0.1001	1	0	12
52.0000	0.7692	0.2308	0.1169	1	0	11
58.0000	0.6923	0.3077	0.1280	1	0	10
60.0000	0.6154	0.3846	0.1349	1	0	9
62.0000	0.5385	0.4615	0.1383	1	0	8
65.0000	0.4615	0.5385	0.1383	1	0	7
67.0000	0.3846	0.6154	0.1349	1	0	6
71.0000	0.2308	0.7692	0.1169	2	0	5
72.0000	0.1538	0.8462	0.1001	1	0	3
73.0000	0.0769	0.9231	0.0739	1	0	2
89.0000	0.0000	1.0000	0.0000	1	0	1

A (continued)

Runx1KO/p21WT

Postnatal day of anagen entrance	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.0000	1.0000	0.0000	0.0000	0	0	10
30.0000	0.9000	0.1000	0.0949	1	0	10
32.0000	0.6000	0.4000	0.1549	3	0	9
34.0000	0.5000	0.5000	0.1581	1	0	6
35.0000	0.3000	0.7000	0.1449	2	0	5
36.0000	0.2000	0.8000	0.1265	1	0	3
38.0000	0.1000	0.9000	0.0949	1	0	2
49.0000	0.0000	1.0000	0.0000	1	0	1

Runx1WT/p21KO

Postnatal day of anagen entrance	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.0000	1.0000	0.0000	0.0000	0	0	21
24.0000	0.9048	0.0952	0.0641	2	0	21
25.0000	0.6190	0.3810	0.1060	6	0	19
26.0000	0.5238	0.4762	0.1090	2	0	13
27.0000	0.3810	0.6190	0.1060	3	0	11
28.0000	0.2857	0.7143	0.0986	2	0	8
29.0000	0.1905	0.8095	0.0857	2	0	6
30.0000	0.1429	0.8571	0.0764	1	0	4
31.0000	0.0952	0.9048	0.0641	1	0	3
32.0000	0.0476	0.9524	0.0465	1	0	2
36.0000	0.0000	1.0000	0.0000	1	0	1

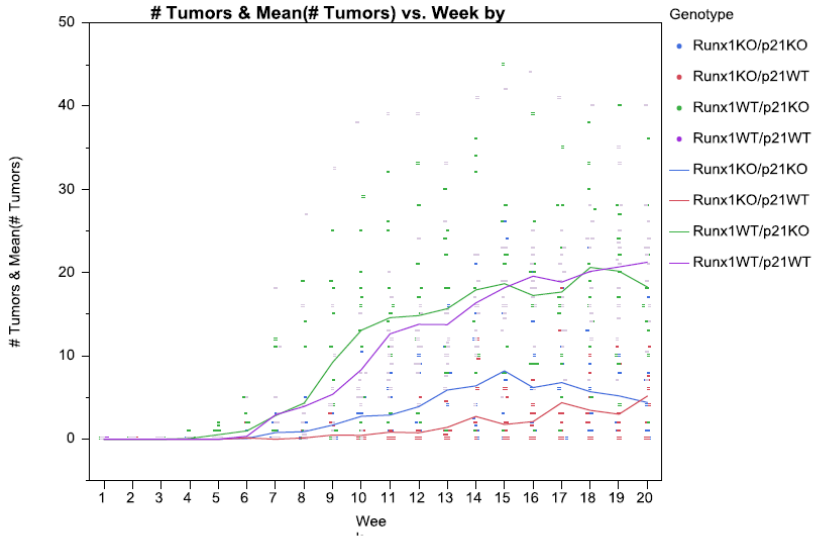
Runx1WT/p21WT

Postnatal day of anagen entrance	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.0000	1.0000	0.0000	0.0000	0	0	15
24.0000	0.9333	0.0667	0.0644	1	0	15
25.0000	0.8667	0.1333	0.0878	1	0	14
28.0000	0.7333	0.2667	0.1142	2	0	13
30.0000	0.5333	0.4667	0.1288	3	0	11
31.0000	0.2000	0.8000	0.1033	5	0	8
32.0000	0.0667	0.9333	0.0644	2	0	3
37.0000	0.0000	1.0000	0.0000	1	0	1

Combined

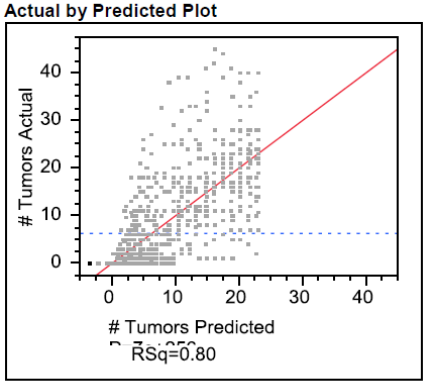
Postnatal day of anagen entrance	Survival	Failure	SurvStdErr	Number failed	Number censored	At Risk
0.0000	1.0000	0.0000	0.0000	0	0	59
24.0000	0.9492	0.0508	0.0286	3	0	59
25.0000	0.8305	0.1695	0.0488	7	0	56
26.0000	0.7966	0.2034	0.0524	2	0	49
27.0000	0.7458	0.2542	0.0567	3	0	47
28.0000	0.6780	0.3220	0.0608	4	0	44
29.0000	0.6441	0.3559	0.0623	2	0	40
30.0000	0.5593	0.4407	0.0646	5	0	38
31.0000	0.4576	0.5424	0.0649	6	0	33
32.0000	0.3559	0.6441	0.0623	6	0	27
34.0000	0.3390	0.6610	0.0616	1	0	21
35.0000	0.3051	0.6949	0.0599	2	0	20
36.0000	0.2712	0.7288	0.0579	2	0	18
37.0000	0.2542	0.7458	0.0567	1	0	16
38.0000	0.2373	0.7627	0.0554	1	0	15
45.0000	0.2203	0.7797	0.0540	1	0	14
48.0000	0.2034	0.7966	0.0524	1	0	13
49.0000	0.1864	0.8136	0.0507	1	0	12
52.0000	0.1695	0.8305	0.0488	1	0	11
58.0000	0.1525	0.8475	0.0468	1	0	10
60.0000	0.1356	0.8644	0.0446	1	0	9
62.0000	0.1186	0.8814	0.0421	1	0	8
65.0000	0.1017	0.8983	0.0393	1	0	7
67.0000	0.0847	0.9153	0.0363	1	0	6
71.0000	0.0508	0.9492	0.0286	2	0	5
72.0000	0.0339	0.9661	0.0236	1	0	3
73.0000	0.0169	0.9831	0.0168	1	0	2
89.0000	0.0000	1.0000	0.0000	1	0	1

B



Each error bar is constructed using a 95% confidence interval of the mean.

Response # Tumors
Whole Model
Actual by Predicted Plot



Summary of Fit

RSquare	0.79772
RSquare Adj	0.796263
Root Mean Square Error	4.266348
Mean of Response	6.322959
Observations (or Sum Wgts)	980

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	-2.700571	0.680702	60.07	-3.97	0.0002*
Genotype[Runx1KO/p21KO]	-3.054251	1.169399	45	-2.61	0.0122*
Genotype[Runx1KO/p21WT]	-4.798674	1.069772	45	-4.49	<.0001*
Genotype[Runx1WT/p21KO]	4.1907486	1.098353	45	3.82	0.0004*
Week	0.8426021	0.02382	927	35.37	<.0001*
Genotype[Runx1KO/p21KO]*(Week-10.5)	-0.423091	0.044	927	-9.62	<.0001*
Genotype[Runx1KO/p21WT]*(Week-10.5)	-0.599716	0.040251	927	-14.90	<.0001*
Genotype[Runx1WT/p21KO]*(Week-10.5)	0.4738139	0.041326	927	11.47	<.0001*

B (continued)

Source	Nparm	DF	DFDen	F Ratio	Prob > F
Genotype	3	3	45	13.6138	<.0001*
Week	1	1	927	1251.262	<.0001*
Genotype*Week	3	3	927	162.4863	<.0001*

LSMeans (Least Square Means) Differences Tukey HSD
α=
0.050
LSMean[i] By LSMean[j]

Mean[i]-Mean[j] Std Err Dif Lower CL Dif Upper CL Dif	Runx1KO/p21KO	Runx1KO/p21WT	Runx1WT/p21KO	Runx1WT/p21WT
Runx1KO/p21KO	0	1.74442	-7.245	-6.7164
	0	1.84949	1.8827	1.82055
	0	-3.1895	-12.267	-11.573
	0	6.67831	-2.2225	-1.8598
Runx1KO/p21WT	-1.7444	0	-8.9894	-8.4609
	1.84949	0	1.76022	1.69358
	-6.6783	0	-13.685	-12.979
	3.18947	0	-4.2937	-3.9429
Runx1WT/p21KO	7.245	8.98942	0	0.52857
	1.8827	1.76022	0	1.72979
	2.22253	4.29368	0	-4.086
	12.2675	13.6852	0	5.14312
Runx1WT/p21WT	6.71643	8.46085	-0.5286	0
	1.82055	1.69358	1.72979	0
	1.85976	3.94288	-5.1431	0
	11.5731	12.9788	4.08598	0

Level			Least Sq Mean
Runx1WT/p21KO	A		10.337500
Runx1WT/p21WT	A		9.808929
Runx1KO/p21KO		B	3.092500
Runx1KO/p21WT		B	1.348077

Contrast
Test Detail

Runx1KO/p21KO	1	1	1
Runx1KO/p21WT	-1	0	0
Runx1WT/p21KO	0	-1	0
Runx1WT/p21WT	0	0	-1
Estimate	1.7444	-7.245	-6.716
Std Error	1.8495	1.8827	1.8205
t Ratio	0.9432	-3.848	-3.689
Prob>|t|	0.3506	0.0004	0.0006

NumDF	DenDF	F Ratio	Prob > F
3	45	13.6138	<.0001*

Contrast
Test Detail

Runx1KO/p21KO,1	0
Runx1KO/p21KO,2	0
Runx1KO/p21KO,3	0
Runx1KO/p21KO,4	0
Runx1KO/p21KO,5	0
Runx1KO/p21KO,6	0
Runx1KO/p21KO,7	0
Runx1KO/p21KO,8	0
Runx1KO/p21KO,9	0
Runx1KO/p21KO,10	0
Runx1KO/p21KO,11	0

B (continued)

Runx1KO/p21KO,12	0
Runx1KO/p21KO,13	**0.25**
Runx1KO/p21KO,14	**0.25**
Runx1KO/p21KO,15	**0.25**
Runx1KO/p21KO,16	**0.25**
Runx1KO/p21KO,17	0
Runx1KO/p21KO,18	0
Runx1KO/p21KO,19	0
Runx1KO/p21KO,20	0
Runx1KO/p21WT,1	0
Runx1KO/p21WT,2	0
Runx1KO/p21WT,3	0
Runx1KO/p21WT,4	0
Runx1KO/p21WT,5	0
Runx1KO/p21WT,6	0
Runx1KO/p21WT,7	0
Runx1KO/p21WT,8	0
Runx1KO/p21WT,9	0
Runx1KO/p21WT,10	0
Runx1KO/p21WT,11	0
Runx1KO/p21WT,12	0
Runx1KO/p21WT,13	**-0.25**
Runx1KO/p21WT,14	**-0.25**
Runx1KO/p21WT,15	**-0.25**
Runx1KO/p21WT,16	**-0.25**
Runx1KO/p21WT,17	0
Runx1KO/p21WT,18	0
Runx1KO/p21WT,19	0
Runx1KO/p21WT,20	0
Runx1WT/p21KO,1	0
Runx1WT/p21KO,2	0
Runx1WT/p21KO,3	0
Runx1WT/p21KO,4	0
Runx1WT/p21KO,5	0
Runx1WT/p21KO,6	0
Runx1WT/p21KO,7	0
Runx1WT/p21KO,8	0
Runx1WT/p21KO,9	0
Runx1WT/p21KO,10	0
Runx1WT/p21KO,11	0
Runx1WT/p21KO,12	0
Runx1WT/p21KO,13	0
Runx1WT/p21KO,14	0
Runx1WT/p21KO,15	0
Runx1WT/p21KO,16	0
Runx1WT/p21KO,17	0
Runx1WT/p21KO,18	0
Runx1WT/p21KO,19	0
Runx1WT/p21KO,20	0
Runx1WT/p21WT,1	0
Runx1WT/p21WT,2	0
Runx1WT/p21WT,3	0
Runx1WT/p21WT,4	0
Runx1WT/p21WT,5	0
Runx1WT/p21WT,6	0
Runx1WT/p21WT,7	0
Runx1WT/p21WT,8	0
Runx1WT/p21WT,9	0
Runx1WT/p21WT,10	0
Runx1WT/p21WT,11	0
Runx1WT/p21WT,12	0
Runx1WT/p21WT,13	0
Runx1WT/p21WT,14	0
Runx1WT/p21WT,15	0
Runx1WT/p21WT,16	0
Runx1WT/p21WT,17	0
Runx1WT/p21WT,18	0
Runx1WT/p21WT,19	0
Runx1WT/p21WT,20	0
Estimate	4.6654
Std Error	1.9999
t Ratio	2.3328
Prob>|t|	0.0229

NumDF	DenDF	F Ratio	Prob > F
1	61.43	5.4422	0.0229*

Figure 3.11. Statistical and Theoretical Methods

Statistical and Theoretical Methods

Much of the experimental interpretation in the main text rests on the quantitative determination for each experimental condition $\mathcal{E}$ (specified by a mouse genotype and the labeling and chase times) of $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$, the mean proportion $(N^{\mathcal{E}}+2)$-vector, having components $\bar{\pi}^{\mathcal{E}}_n$ $(0 \le n \le N^{\mathcal{E}})$, which are the fractions of labeled bulge cells that have replicated n times, and $\bar{\pi}^{\mathcal{E}}_{N^{\mathcal{E}}+1}$, which is the fraction that has divided more than $N^{\mathcal{E}}$ times. As explained in the main text, this is based on the equivalence between the number of cell divisions and the number of two-fold dilutions of the cellular H2B-GFP fluorescence. For each $\mathcal{E}$, the bulge cell FACS data $\mathcal{X}^m$ from mouse replicate m $(1 \le m \le M)$ in combination with background subtraction data $\mathcal{X}^{\mathcal{U}}$, which was used to adjust for the fraction of non-fluorescent unlabeled cells, was analyzed to estimate the posterior distribution $p(\boldsymbol{\pi}^m|\mathcal{X}^m,\mathcal{X}^{\mathcal{U}})$ of the $\boldsymbol{\pi}^m$, the replication proportion (N^m+2)-vector for mouse m alone. (Because of differences in FACS resolution, in some cases $N^m < N^{\mathcal{E}}$.) The M posteriors were then combined using a hierarchical Bayesian analysis to compute $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\{\mathcal{X}^m\},\mathcal{X}^{\mathcal{U}})$, where $\{\mathcal{X}^m\}$ is the set of all M mouse data sets.

For example, Figure S2A shows the histogram of the H2B-GFP FACS data obtained from the FITC channel of the live bulge cells from one mouse. In this case, $N^m = 5$ because we are able to distinguish five well-formed but overlapping (roughly) Gaussian peaks and the upper half of a semi-resolved additional "lowest discernible peak" (LDP). Peak $n = 0$, which has the highest fluorescence, represents undivided cells, $n = 1$ represents cells that have divided one time, and the LDP, $n = 5$, represents cells that have divided five times. We call these "division-resolved" cells. Cells with less fluorescence represent a mixture of "highly replicated" cells that have divided six or more times $(n = 6)$ and unlabeled cells that never expressed H2B-GFP, presumably (1) because of mosaicism in the transgenes of the *tet*-repressible pulse-chase system (2). In some cases there were no highly replicated cells and/or the LDP was fully resolved.

Determining $p(\boldsymbol{\pi}^m|\mathcal{X}^m,\mathcal{X}^{\mathcal{U}})$ faced three challenges: 1) Deconvolving the fully resolved but overlapping components $(0 \le n \le 4$ in Figure S2A), 2) deconvolving the semi-resolved LDP, and 3) deconvolving the highly replicated and unlabeled cells in the low-fluorescence region, which in this case lies below the center of the LDP.[1] These were addressed by separating $\mathcal{X}^m$ into two parts—$\mathcal{X}^m_{\mathsf{lo}}$, which contains the low-fluorescence data, and $\mathcal{X}^m_{\mathsf{res}}$, which contains the division-resolved data—and analyzing them separately. $\mathcal{X}^m_{\mathsf{lo}}$, with additional information from $\mathcal{X}^{\mathcal{U}}$—a reference collection of data sets from unchased and short-chased experiments where the low-fluorescence region contained only unlabeled and no highly replicated cells—was used to determine the distribution of the proportion of highly replicated cells. $\mathcal{X}^m_{\mathsf{res}}$ was deconvolved into a Gaussian mixture of $0 \le n \le N^m$ components. A special procedure was used when the LDP was only semi-resolved: In these cases $\mathcal{X}^m_{\mathsf{res}}$ was first reflected about the center of the LDP to generate a symmetric data set having $2N^m + 1$ well-formed peaks.

It remained to combine all M $p(\boldsymbol{\pi}^m|\mathcal{X}^m,\mathcal{X}^{\mathcal{U}})$ to compute $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\{\mathcal{X}^m\},\mathcal{X}^{\mathcal{U}})$. While experimental factors caused variations in the estimated values of the $\boldsymbol{\pi}^m$, the true $\boldsymbol{\pi}^m$ differed from each other because of inherent biological variations between mice from the same inbred strain: each $\boldsymbol{\pi}^m$ was an independent selection from a probability distribution $p(\boldsymbol{\pi}^m|\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}})$, where $\Sigma^{\mathcal{E}}$ is the variance governing the mouse strain biological proportion vector probability distribution. This stochastic process was also incorporated into the analysis.

We used Bayesian inversion of this multistep process to determine $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\{\mathcal{X}^m\},\mathcal{X}^{\mathcal{U}})$ as described in the first section below. It was used to compute the mean values and error estimates for the bar charts shown in Figures 2 and S2B. Its use in determining the average number of divisions and to assess the

---

[1] The low-fluorescence region boundary is set to just below the LDP when it is fully resolved.

Figure 3.11. continued

goodness of fit of two dynamical models of bulge cell proliferation is described in the second section. This showed that cell replication over PD21-47 is consistent with a simple single-population Poisson model for cell proliferation in the WT, but not in the p21 KO, mice. The KO replication pattern was shown to be consistent with a model that incorporated a subpopulation of cells that replicated at least three times more than the WT cells over the entire hair cycle. The possibility that this subpopulation represents a small fraction of cells that respond to p21 KO by increasing their replication rate or by escaping the suppression of cell replication that normally occurs at the end of early anagen is explored in the third section.

I. COMPUTING $p(\pi^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})$

The analysis is based on a hierarchical stochastic model that generates a data set $\mathcal{X}^m$ for each mouse m. $\mathcal{X}^{\mathcal{E}}$ is the collection $\{\mathcal{X}^m\}$ for experimental conditions $\mathcal{E}$. With experimental conditions for which there is little or no labeled cell division, the process also generates $\mathcal{X}^{\mathcal{U}}$, a collection of "background" data sets that can be used to ascertain the *a priori* distribution of f_u, the fraction of unlabeled cells.

The key elements of the model are: 1) a biological mouse selection process, governed by the proportion $(N^{\mathcal{E}}+2)$-vector $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ and $\Sigma^{\mathcal{E}}$, which determine the vectors $\boldsymbol{\pi}^m$ describing the proportion of the different cell replication subpopulations in mouse m, 2) an experimental background process that generates f_u^m, and 3) an experimental process that generates the (N^m+1)-vector variates $\tilde{\boldsymbol{\mu}}^m$ and $\tilde{\boldsymbol{\lambda}}^m$ that describe the positions and widths of the N^m+1 discernible Gaussians in the FACS data histograms. Because of differing experimental resolutions, N^m may differ between mice. To use as much data as possible subject to the requirement that can estimate the variance of each of the $N^{\mathcal{E}}+2$ proportion vector components, we set $N^{\mathcal{E}}$ equal to the largest value for which there are at least two $\mathcal{X}^m$ having $N^m \geq N^{\mathcal{E}}$. Our task is to compute the posterior probability distribution

$$p(\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}) \propto p(\mathcal{X}^{\mathcal{E}}|\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})\, p(\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}}) \tag{1}$$

and then to reduce this to a posterior distribution over $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$. (Here and below unconditional probability distributions are priors.)

We proceed by separating $\mathcal{X}^{\mathcal{E}}$ into subcomponents and identifying sufficient statistics for each. The first step is to factorize $\mathcal{X}^{\mathcal{E}}$: Because the mice are independent and $\boldsymbol{\pi}^m$ is sufficient for $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ and $\Sigma^{\mathcal{E}}$

$$\begin{aligned} p(\mathcal{X}^{\mathcal{E}}|\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}) &= \prod_{m=1}^{M} p(\mathcal{X}^m|\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}) \\ &= \prod_{m=1}^{M} \int p(\mathcal{X}^m|\boldsymbol{\pi}^m,\mathcal{X}^{\mathcal{U}})\, p(\boldsymbol{\pi}^m|\bar{\boldsymbol{\pi}}^{\mathcal{E}},\Sigma^{\mathcal{E}})\, d\boldsymbol{\pi}^m \,. \end{aligned} \tag{2}$$

(Here and below we implicitly assume that integrations over proportion vectors and their components satisfy the constraints

$$\begin{aligned} 0 \leq \pi_n \leq 1 \\ 1 = \boldsymbol{\pi}\cdot\mathbf{1}\,, \end{aligned} \tag{3}$$

where $\cdot$ denotes the inner product and $\mathbf{1}$ is the vector having all elements equal to one and dimensionality set by context; in this case it is the dimensionality of $\boldsymbol{\pi}$. Consistent with this, all covariance matrices satisfy

$$\mathbf{1}\cdot\Sigma = \Sigma\cdot\mathbf{1} = 0\,.) \tag{4}$$

We next separate $\mathcal{X}^m$ into $\mathcal{X}^m_{\text{lo}}$ and $\mathcal{X}^m_{\text{res}}$ as discussed above. Since we do not distinguish between the intensities of the data points in $\mathcal{X}^m_{\text{lo}}$, it is sufficiently described by the fraction of data points that it contains. We deterministically separate this fraction into two parts (Sec. I D): that coming from the LDP and the

Figure 3.11. continued

stochastic remainder, F^m_{amb}, which represents the mixture of unlabeled and highly replicated cells, i.e., the ambiguous fraction. Therefore, restricting to sufficient stochastic dependencies,[2]

$$\mathcal{X}^m \equiv \{F^m_{\mathsf{amb}}, \mathcal{X}^m_{\mathsf{res}}\} \; . \tag{5}$$

F^m_{amb} and $\mathcal{X}^m_{\mathsf{res}}$ are independent.

$\mathcal{X}^m_{\mathsf{res}}$ does not depend on $\pi^m_{N^m+1}$, so it is helpful to separate $\boldsymbol{\pi}^m$ into $\pi^m_{N^m+1}$ and $\tilde{\boldsymbol{\pi}}^m$, the normalized projected proportion (N^m+1)-vector with components proportional to the $\{\pi^m_n : 0 \leq n \leq N^m\}$:

$$\tilde{\pi}^m_n = \frac{\pi^m_n}{1 - \pi^m_{N^m+1}} \qquad (0 \leq n \leq N^m)$$
$$\mathbf{1} \cdot \tilde{\boldsymbol{\pi}}^m = 1 \; .$$

With this notation,

$$p(\mathcal{X}^m_{\mathsf{res}} | \boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}}) = p(\mathcal{X}^m_{\mathsf{res}} | \tilde{\boldsymbol{\pi}}^m) \; . \tag{6}$$

Conversely, F^m_{amb} does not depend on $\tilde{\boldsymbol{\pi}}^m$, so

$$p(F^m_{\mathsf{amb}} | \boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}}) = p(F^m_{\mathsf{amb}} | \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}}) \; . \tag{7}$$

Eqs. (5), (6), and (7) imply that

$$\begin{aligned} p(\mathcal{X}^m | \boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}}) &= p(F^m_{\mathsf{amb}} | \boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}}) \, p(\mathcal{X}^m_{\mathsf{res}} | \boldsymbol{\pi}^m) \\ &= p(F^m_{\mathsf{amb}} | \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}}) \, p(\mathcal{X}^m_{\mathsf{res}} | \tilde{\boldsymbol{\pi}}^m) \; . \end{aligned} \tag{8}$$

Combining Eqs. (1), (2), and (8) gives

$$\begin{aligned} &p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \propto \\ &\quad p(\Sigma^{\mathcal{E}}) \prod_{m=1}^{M} \int p(F^m_{\mathsf{amb}} | \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}}) \times \\ &\quad p(\mathcal{X}^m_{\mathsf{res}} | \tilde{\boldsymbol{\pi}}^m) \, p(\tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1} | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}}) \, d\tilde{\boldsymbol{\pi}}^m \, d\pi^m_{N^m+1} \; , \end{aligned} \tag{9}$$

where we assume that $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}}) = p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}) \, p(\Sigma^{\mathcal{E}})$ and that $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}})$ is non-informative.

Applying Bayes' Theorem,[3]

$$p(\mathcal{X}^m_{\mathsf{res}} | \tilde{\boldsymbol{\pi}}^m) = \frac{p(\tilde{\boldsymbol{\pi}}^m | \mathcal{X}^m_{\mathsf{res}}) \, p(\mathcal{X}^m_{\mathsf{res}})}{p(\tilde{\boldsymbol{\pi}}^m)}$$

gives

$$\begin{aligned} &p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \propto p(\Sigma^{\mathcal{E}}) \times \\ &\quad \prod_{m=1}^{M} \int p(F^m_{\mathsf{amb}} | \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}}) \times \\ &\quad \frac{p(\tilde{\boldsymbol{\pi}}^m | \mathcal{X}^m_{\mathsf{res}})}{p(\tilde{\boldsymbol{\pi}}^m)} \, p(\tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1} | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}}) \, d\tilde{\boldsymbol{\pi}}^m \, d\pi^m_{N^m+1} \; . \end{aligned} \tag{10}$$

---

[2] We analyze $\mathcal{X}^m$ conditional on the number of labeled cells, so the number is a parameter, not a stochastic variate, and will not be explicitly denoted. Although the analysis is also conditional on $\mathcal{X}^{\mathcal{U}}$, we do explicitly denote it to emphasize its role in determining the posterior distribution over F^m_{amb} as in Eq. (7).

[3] In principle, $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$ could be computed directly using the likelihoods $\{p(\mathcal{X}^m_{\mathsf{res}} | \tilde{\boldsymbol{\pi}}^m)\}$. However, $\mathcal{X}^m_{\mathsf{res}}$ has a complicated structure that depends on, in addition to $\tilde{\pi}^m$, the variates $\{\tilde{\mu}^m, \tilde{\lambda}^m\}$, which determine the Gaussian centers and widths (see Sec. I C). The required marginalizations over these variates would require unfeasible multidimensional integrations. The advantage of expressing Eq. (9) using the posterior distribution $p(\tilde{\pi}^m | \mathcal{X}^m_{\mathsf{res}}, \mathcal{X}^{\mathcal{U}})$ is that it can be approximately factorized (Sec. I C), thereby making the marginalizations trivial.

Figure 3.11. continued

We use Eq. (10) to compute the posterior distribution over $(\tilde{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}})$. We describe the dissection of $\mathcal{X}^m$ in Sec. I B, the computation of $p(\tilde{\boldsymbol{\pi}}^m|\mathcal{X}^m_{\mathsf{res}})$ and $p(F^m_{\mathsf{amb}}|\pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}})$ in Secs. I C, I D and I E, the marginalization over the $\{\tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1}\}$ in Sec. I F, and the maximum likelihood treatment of $\Sigma^{\mathcal{E}}$ in Sec. I G. The next section describes the first computational step: identifying the live bulge cell data.

A. Identifying the live bulge cells in the FACS multichannel data

The live bulge cells were identified in the mouse epidermis preparations by appropriate gating of the multichannel FACS data: Manual gating of raw FACS data isolated single cells (based on forward and side scatter; Figure S1A, first two panels) that were live (i.e., propidium iodide negative detected in the PerCp-Cy5.5 channel; third panel) and resided within the bulge (as indicated by bulge markers α6-integrin and CD34 detected in the PE and APC channels, respectively; fourth panel).

B. Dissecting $\mathcal{X}^m$

$\mathcal{X}^m$ consists of the set of $X^m \sim 50,000$ H2B-GFP fluorescence log-intensity[4] measurements $\{x_i : 1 \le i \le X^m\}$.[5] We decompose it into $\mathcal{X}^m_{\mathsf{lo}}$, the data lying below a lower log-intensity cutoff x^m_{lo}, and the division-resolved data $\mathcal{X}^m_{\mathsf{res}}$, which can be deconvolved into a mixture of Gaussians (possibly with a half-Gaussian semi-resolved LDP). When the LDP is fully resolved (i.e., in the short-chase PD24–27 and PD36–42 data sets), x^m_{lo} is subjectively set to its left boundary with little ambiguity. When it is only semi-resolved (i.e., in the long-chase PD21–35 and PD21–47 data sets), x^m_{lo} is algorithmically set to the center of the LDP.

For this purpose we use Gaussian kernel density smoothing to estimate the probability density $p(x)$ corresponding to the set $\{x_i : 1 \le i \le X^m\}$ of fluorescence log-intensity measurements x_i in $\mathcal{X}^m$:

$$p(x) = \frac{1}{X^m} \sum_{i=1}^{X^m} \mathcal{N}(x; x_i, h^2) \,,$$

where

$$\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$$

is the normal distribution and the bandwidth h determines the degree of smoothing and hence, the number of apparent peaks $K(h)$ in $p(x)$. h^*, the appropriate value of h, is the smallest value that has $K(h) = N^m + 1$ peaks. $K(h)$ is defined as the number of local minima of the second derivative of $p(x)$, which are separated by second derivative zero-crossings. This use of a second (rather than first) derivative condition enables us to detect incomplete "peaks" that only form shoulders in the distribution (e.g., the LDP), even though they are not local maxima. To exclude peaks within the noisy low-fluorescence region, the examination is limited to x above a subjectively chosen cutoff that is just below the center of the LDP. Once h^* has been determined by one-dimensional search, x^m_{lo} is computed as the lowest minimizer of $d^2p(x)/dx^2$ in the examined region. $\mathcal{X}^m_{\mathsf{lo}}$ is then defined as the subset of data lying below x^m_{lo},

$$\mathcal{X}^m_{\mathsf{lo}} \equiv \{x_i : x_i < x^m_{\mathsf{lo}}\}$$

and F^m_{lo} is defined as the fraction of data included in this subset. Conversely

$$\mathcal{X}^m_{\mathsf{res}} = \{x_i : x^m_{\mathsf{lo}} \le x_i\}$$

---

[4] Although the logicle transformation (3) is often applied to FACS fluorescence intensities because of its ability to handle negative values that result from FACS background subtractions, for our purposes it is simpler and more efficient to simply assign all negative values to the smallest positive intensity and to use logarithmic variables.

[5] The FACS aggregates the $\sim 50,000$ x_i into $\sim 12,500$ unique bins, which permits expedited computation. For simplicity we ignore this in the discussion.

Figure 3.11. continued

and F_{res}^m is defined as the fraction of data in this subset.

The dissection satisfies

$$1 = F_{\mathsf{lo}}^m + F_{\mathsf{res}}^m \,.$$

C. $p(\tilde{\pi}^m|\mathcal{X}_{\mathsf{res}}^m)$

To simplify notation, in this section we suppress m superscripts.

1. Preprocessing

We model $\mathcal{X}_{\mathsf{res}}$ as generated by a stochastic Gaussian mixture process parameterized by the $(N+1)$-vectors $\tilde{\boldsymbol{\pi}}$, $\tilde{\boldsymbol{\mu}}$, and $\tilde{\boldsymbol{\lambda}}$, the proportions, mean log-intensities, and precisions of the Gaussian components. We use the mean-field variational Bayes approach, which expedites computation while providing a Bayesian posterior distribution, not just a maximum likelihood estimate (4). However, it is not possible to directly apply the method to data sets that have non-Gaussian distortions coming from a semi-resolved LDP and/or an $n = 0$ peak with an overly extended right tail. Instead, they are preprocessed to make them amenable to the variational method. The posterior distributions computed from these data sets are subsequently postprocessed (Sec. I C 5) to compensate.

In cases where the extended right tail of peak $n = 0$ has enough mass to distort the Gaussian mixture modeling, an upper cutoff $x_>$, located near the upper base of the peak, is manually selected and the data above, comprising the fraction $f_>$ of F_{res} is removed.

In addition, if the LDP is semi-resolved we prepare a symmetrized data set by reflecting $\mathcal{X}_{\mathsf{res}}$ about x_{lo}, which in these cases is at the center of the LDP (Sec. I B). This data set has $2N+1$ Gaussian components.

The data set produced after neither, either, or both modifications is designated $\widetilde{\mathcal{X}}$, which we model using $\widetilde{N}+1$ Gaussians. $\widetilde{N} = N$ if the data set was not reflected (i.e., fully resolved LDP) and $\widetilde{N} = 2N$ if it was; coordinately, the number of data points was $\widetilde{X} = X$ or $\widetilde{X} = 2X$.

2. Variational Bayesian Gaussian mixture modeling

This method is described in detail in Bishop (4). We provide a brief summary of the salient features.

The Gaussian mixture process generates each data point x_i in two steps: First, a categorical process governed by $\tilde{\boldsymbol{\pi}}$ selects Gaussian n according to

$$p(\boldsymbol{z}^i|\tilde{\boldsymbol{\pi}}) = \prod_{n=0}^{\widetilde{N}} \tilde{\pi}_n^{z_n^i} \qquad (1 \le i \le \widetilde{X}) \,, \tag{11}$$

where the latent variate $\boldsymbol{z}^i$ is a unit $(\widetilde{N}+1)$-vector having $z_n^i = 1$ if Gaussian n is responsible for data point i and $z_n^i = 0$, otherwise. (We denote the complete set as $\mathcal{Z} \equiv \{\boldsymbol{z}^i : 1 \le i \le \widetilde{X}\}$.) Second, the value of x_i is generated by a normally distributed process governed by $\tilde{\boldsymbol{\mu}} \cdot \boldsymbol{z}^i$ and $\tilde{\boldsymbol{\lambda}} \cdot \boldsymbol{z}^i$, the mean and precision of the selected Gaussian:

$$p(x_i|\boldsymbol{z}^i, \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\lambda}}) = \mathcal{N}(x_i; \tilde{\boldsymbol{\mu}} \cdot \boldsymbol{z}^i, 1/\tilde{\boldsymbol{\lambda}} \cdot \boldsymbol{z}^i) \,. \tag{12}$$

Taking the product of Eqs. (11) and (12) over i gives

$$\begin{aligned} p(\mathcal{Z}|\tilde{\boldsymbol{\pi}}) &= \prod_{n=0}^{\widetilde{N}} \prod_{i=1}^{\widetilde{X}} \tilde{\pi}_n^{z_n^i} \\ &= \prod_{n=0}^{\widetilde{N}} \tilde{\pi}_n^{Z_n} \end{aligned} \tag{13}$$

Figure 3.11. continued

and

$$p(\tilde{\mathcal{X}}|\mathcal{Z},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}) = \prod_{n=0}^{\tilde{N}}\prod_{i=1}^{\tilde{X}} \mathcal{N}(x_i;\tilde{\mu}_n,1/\tilde{\lambda}_n)^{z_n^i}$$
$$= \frac{1}{Z_n}\prod_{n=0}^{\tilde{N}}\left(\frac{\tilde{\lambda}_n}{2\pi}\right)^{(Z_n-1)/2} \mathcal{N}(\bar{x}_n;\tilde{\mu}_n,1/\tilde{\lambda}_n Z_n)\, e^{-S_n\tilde{\lambda}_n Z_n/2}\,, \tag{14}$$

where we have defined three statistics of $\mathcal{Z}$:

$$Z_n = \sum_{i=1}^{\tilde{X}} z_n^i \tag{15a}$$

$$\bar{x}_n = \frac{1}{Z_n}\sum_{i=1}^{\tilde{X}} z_n^i\, x_i \tag{15b}$$

$$S_n = \frac{1}{Z_n}\sum_{i=1}^{\tilde{X}} z_n^i\,(x_i-\bar{x}_n)^2\,. \tag{15c}$$

Therefore, the complete likelihood is

$$p(\tilde{\mathcal{X}},\mathcal{Z}|\tilde{\boldsymbol{\pi}},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}) = p(\tilde{\mathcal{X}}|\mathcal{Z},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}})\,p(\mathcal{Z}|\tilde{\boldsymbol{\pi}})\,.$$

For analytic simplicity we choose the prior distribution to be conjugate to this. It is the product of the conjugate priors to $p(\tilde{\mathcal{X}}|\mathcal{Z},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}})$ and $p(\mathcal{Z}|\tilde{\boldsymbol{\pi}})$:

$$p(\tilde{\boldsymbol{\pi}},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}) = p(\tilde{\boldsymbol{\pi}})\prod_{n=0}^{\tilde{N}} p(\tilde{\mu}_n|\tilde{\lambda}_n)p(\tilde{\lambda}_n)\,; \tag{16}$$

where

$$p(\tilde{\boldsymbol{\pi}}) = \mathcal{D}(\tilde{\boldsymbol{\pi}};\{\alpha_n^0\}) \tag{17a}$$
$$p(\tilde{\mu}_n|\tilde{\lambda}_n) = \mathcal{N}[\tilde{\mu}_n;\tilde{\mu}_n^0,1/(\beta_n^0\tilde{\lambda}_n)] \tag{17b}$$
$$p(\tilde{\lambda}_n) = \mathcal{G}(\tilde{\lambda}_n;a_n^0,b_n^0)\,; \tag{17c}$$

$$\mathcal{D}(\boldsymbol{\pi};\{\alpha_n\}) = \frac{\Gamma(\alpha_\Sigma)}{\prod_{n=0}^{\tilde{N}}\Gamma(\alpha_n)}\prod_{n=0}^{\tilde{N}}\pi_n^{\alpha_n-1} \tag{18}$$
$$\alpha_\Sigma \equiv \sum_{n=0}^{\tilde{N}}\alpha_n$$

is the Dirichlet distribution; and

$$\mathcal{G}(x;a,b) = \frac{b^a x^{a-1}}{\Gamma(a)}e^{-bx} \qquad\qquad (0\le x<\infty)$$

is the gamma distribution. $\Gamma(\cdot)$ is the gamma function.[6]

---

[6] Eq. (17a) follows from the observation that $p(\mathcal{Z}|\tilde{\boldsymbol{\pi}})$ is proportional to a multinomial distribution whose conjugate prior is the Dirichlet distribution. Eqs. (17b) and (17c) follow from the observation that $p(\tilde{\mathcal{X}}|\mathcal{Z},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}})$ is proportional to a Gaussian-gamma distribution, which is self-conjugate.

Figure 3.11. continued

$\mathcal{D}(\boldsymbol{\pi};\{\alpha_n\})$ has mean and covariance matrix

$$\bar{\pi}_n = \alpha_n/\alpha_\Sigma \qquad (0 \le n \le \tilde{N}) \tag{19a}$$

$$\Sigma^{\mathcal{D}}_{nn'} = \frac{\alpha_n\,(\delta_{nn'}\,\alpha_\Sigma - \alpha_{n'})}{(\alpha_\Sigma)^2(\alpha_\Sigma+1)} \qquad (0 \le n,n' \le \tilde{N})\,. \tag{19b}$$

$\mathcal{G}(x;a,b)$ has mode and variance

$$\mathrm{Mo}\,x = (a-1)/b \qquad (a \ge 1) \tag{20a}$$

$$\mathrm{Var}\,x = a/b^2\,. \tag{20b}$$

Calculation of the posterior distribution $p(\mathcal{Z},\tilde{\boldsymbol{\pi}},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}|\widetilde{\mathcal{X}})$ is complicated by the coupling of the latent variables $\mathcal{Z}$ to the parameters $\tilde{\boldsymbol{\pi}}$, $\tilde{\boldsymbol{\mu}}$, and $\tilde{\boldsymbol{\lambda}}$. The variational Bayesian approach addresses this difficulty by assuming an approximate, factorized posterior distribution

$$p(\mathcal{Z},\tilde{\boldsymbol{\pi}},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}|\widetilde{\mathcal{X}}) \approx p(\mathcal{Z}|\widetilde{\mathcal{X}})\,p(\tilde{\boldsymbol{\pi}},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}|\widetilde{\mathcal{X}})\,. \tag{21}$$

The "best" form of this factorization is determined by minimizing the Kullback-Liebler divergence between the left- and right-hand sides. This leads to a "mean-field" approximation in which the conditional distributions are decoupled. Without additional assumption, Eqs. (16), (17), and (21) imply that

$$p(\tilde{\boldsymbol{\pi}},\tilde{\boldsymbol{\mu}},\tilde{\boldsymbol{\lambda}}|\widetilde{\mathcal{X}}) = p(\tilde{\boldsymbol{\pi}}|\widetilde{\mathcal{X}})\prod_{n=0}^{\tilde{N}} p(\tilde{\mu}_n|\tilde{\lambda}_n,\widetilde{\mathcal{X}})p(\tilde{\lambda}_n|\widetilde{\mathcal{X}})\,,$$

where

$$p(\tilde{\boldsymbol{\pi}}|\widetilde{\mathcal{X}}) = \mathcal{D}(\tilde{\boldsymbol{\pi}};\{\alpha_n\}) \tag{22a}$$

$$p(\tilde{\mu}_n|\tilde{\lambda}_n,\widetilde{\mathcal{X}}) = \mathcal{N}[\tilde{\mu}_n;\bar{\tilde{\mu}}_n,1/(\beta_n\tilde{\lambda}_n)] \tag{22b}$$

$$p(\tilde{\lambda}_n|\widetilde{\mathcal{X}}) = \mathcal{G}(\tilde{\lambda}_n;a_n,b_n)\,. \tag{22c}$$

Minimization of the Kullback-Liebler divergence via iteration leads to the self-consistent solution for the hyperparameters[7]

$$\alpha_n = \alpha^0_n + \hat{Z}_n \tag{23a}$$

$$\bar{\tilde{\mu}}_n = (\beta^0_n\bar{\tilde{\mu}}^0_n + \hat{\bar{x}}_n\hat{Z}_n)/\beta_n \tag{23b}$$

$$\beta_n = \beta^0_n + \hat{Z}_n \tag{23c}$$

$$a_n = a^0_n + \hat{Z}_n/2 \tag{23d}$$

$$b_n = b^0_n + \hat{S}_n\hat{Z}_n/2 + \frac{\beta^0_n\hat{Z}_n}{2\beta_n}\,(\hat{\bar{x}}_n - \bar{\tilde{\mu}}^0_n)^2\,, \tag{23e}$$

where $\hat{Z}_n$, $\hat{\bar{x}}_n$, and $\hat{S}_n$ are defined analogously to Eqs. (15) with the z^i_n replaced by their means, the responsibilities

$$r^i_n \equiv E[z^i_n]\,.$$

3. Setting the prior hyperparameters

We want the prior distributions to be as uninformative as possible. However, the Kullback-Liebler divergence has multiple local minima over the space of posterior hyperparameters and it is necessary to guide

[7] Readers should be aware that at least some versions of Bishop (4) contain a typographical error in this equation.

Figure 3.11. continued

the solution to the correct catchment region by using a weakly constraining prior. A common problem is improper convergence to minima corresponding to overly-wide Gaussians that span two or more peaks. This can be avoided by mildly constraining the $\tilde{\boldsymbol{\lambda}}$ prior. Subjective examination of the FACS histograms suggests that a reasonable estimate of the standard deviation of each Gaussian (i.e., $1/\sqrt{\tilde{\lambda}_n}$) is about one-third of the separation from its neighbors. Since the expected two-fold dilution in H2B-GFP fluorescence with each cell division implies that the Gaussians will be separated by $\Delta\mu \approx \log 2$, we set the prior mode Mo^0 of the precision to

$$\mathrm{Mo}^0\, \tilde{\lambda}_n = (3/\log 2)^2 \qquad (\forall\, n)\ .$$

We found that constraining the standard deviation of $\tilde{\boldsymbol{\lambda}}$ to roughly one-fifth this value was adequate for convergence, so we used the additional condition on the prior variance Var^0

$$\sqrt{\mathrm{Var}^0\, \tilde{\lambda}_n} = \mathrm{Mo}^0\, \tilde{\lambda}_n/5 \qquad (\forall\, n)\ .$$

Combining these conditions with Eqs. (20) gives $(a_n^0, b_n^0) \approx (27, 1.4)$, $\forall\, n$.

We found that we could make the priors over $\tilde{\mu}_n$ almost uninformative while maintaining correct convergence: $\bar{\tilde{\mu}}_n^0$ is set to the peak position identified by the kernel density smoothing algorithm (Sec. I B), but is only weakly constrained by setting $\beta_n^0 = 0.003$. This small value results in the mode of the standard deviation of $\tilde{\mu}_n$ being $(\mathrm{Mo}^0\, \sigma_\mu)/\Delta\mu = (\Delta\mu \sqrt{\beta^0\ \mathrm{Mo}^0\, \tilde{\lambda}})^{-1} = 1/(3\sqrt{0.003}) \sim 6$ times the inter-Gaussian spacing $\Delta\mu$.

We were able to maintain proper convergence using a completely uninformative prior for $\tilde{\boldsymbol{\pi}}$:

$$\begin{aligned} \alpha_n^0 &= 1 \qquad (\forall\, n) \\ p(\tilde{\boldsymbol{\pi}}) &= \tilde{N}!\ . \end{aligned} \tag{24}$$

4. The predictive density and the posterior distribution over $\tilde{\boldsymbol{\pi}}$

Eqs. (11) and (12) imply that

$$\begin{aligned} p(x_i|\tilde{\boldsymbol{\pi}}, \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\lambda}}) &= \sum_{\boldsymbol{z}^i} p(x_i|\boldsymbol{z}^i, \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\lambda}})\, p(\boldsymbol{z}^i|\tilde{\boldsymbol{\pi}}) \\ &= \sum_{n=0}^{\tilde{N}} \tilde{\pi}_n \mathcal{N}(x_i; \tilde{\mu}_n, 1/\tilde{\lambda}_n)\ , \end{aligned}$$

where the sum in the first line is over the $\tilde{N}+1$ possible values of $\boldsymbol{z}^i$. Therefore, $p(x|\widetilde{\mathcal{X}})$, the predictive density (i.e., the probability density that a new data point will have log-intensity x), is

$$\begin{aligned} p(x|\widetilde{\mathcal{X}}) &= \int p(x|\tilde{\boldsymbol{\pi}}, \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\lambda}})\, p(\tilde{\boldsymbol{\pi}}, \tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\lambda}}|\widetilde{\mathcal{X}})\, d\tilde{\boldsymbol{\pi}}\, d\tilde{\boldsymbol{\mu}}\, d\tilde{\boldsymbol{\lambda}} \\ &\approx \sum_{n=0}^{\tilde{N}} \int \tilde{\pi}_n\, \mathcal{D}(\tilde{\boldsymbol{\pi}}; \{\alpha_n\})\, d\tilde{\boldsymbol{\pi}} \times \\ &\quad \int \mathcal{N}(x; \tilde{\mu}_n, 1/\tilde{\lambda}_n)\ \mathcal{N}[\tilde{\mu}_n; \bar{\tilde{\mu}}_n, 1/(\beta_n \tilde{\lambda}_n)] \times \\ &\quad \mathcal{G}(\tilde{\lambda}_n; a_n, b_n)\, d\tilde{\mu}_n\, d\tilde{\lambda}_n \\ &= \sum_{n=0}^{\tilde{N}} \frac{\alpha_n}{\alpha_\Sigma} \mathcal{S}t[x; \bar{\tilde{\mu}}_n, (a_n/b_n)\, \beta_n/(1+\beta_n), 2a_n]\ , \end{aligned} \tag{25}$$

Figure 3.11. continued

where the non-standardized Student's t-distribution, which arises from convolving the normal distributions over $\tilde{\mu}_n$ and compounding the resultant normal distribution with the Gamma distribution over $\tilde{\lambda}_n$, is

$$\mathcal{S}t(x|\mu,\lambda,\nu) = \frac{\Gamma(\nu/2+1/2)}{\Gamma(\nu/2)}\left(\frac{\lambda}{\pi\nu}\right)^{1/2} \times \left[1+\frac{\lambda\,(x-\mu)^2}{\nu}\right]^{-\nu/2-1/2} .$$

$p(x|\widetilde{\boldsymbol{X}})$ was used to generate the mixture model curves displayed in Figures 2C, S1A, and S3A,B,C. (Because $\tilde{X} \sim 50,000$, even the smallest $\hat{Z}_n$ is large. Therefore the number of degrees-of-freedom in the Student's t-distribution, $\nu = 2a_n = 2a_n^0 + \hat{Z}_n$, is large and each peak is very close to Gaussian.)

5. Postprocessing

a. Semi-resolved LDP: If $\widetilde{\boldsymbol{X}}$ is a reflected data set, the posterior distributions $p(\tilde{\mu}_n|\tilde{\lambda}_n,\widetilde{\boldsymbol{X}})$ and $p(\tilde{\lambda}_n|\widetilde{\boldsymbol{X}})$ for the extra components with $N < n \leq 2N+1$ will be duplicates of those for the components with $0 \leq n < N$, so we can simply drop them. Since both the components and the associated data points were reflected, the number of data points per component is unchanged and all the data points associated with a component are independent. Therefore, $p(\tilde{\mu}_n|\tilde{\lambda}_n,\widetilde{\boldsymbol{X}})$ and $p(\tilde{\lambda}_n|\widetilde{\boldsymbol{X}})$ are correct without modification for $0 \leq n < N$. However, the reflection doubles the points used in the analysis of the single LDP, $n = N$, component and only half are independent. This incorrectly decreases the apparent variance of the LDP posterior distributions and requires adjustment. It is easy to adjust these distributions to remove this artifact by replacing the $\hat{Z}_n$ in Eqs. (23)b–e by $\hat{Z}_n^S = (1-\delta_{nN}/2)\hat{Z}_n$. We denote the adjusted parameters as $\bar{\tilde{\mu}}_n{}^S$, β_n^S, a_n^S, and b_n^S and use them in Eq. (25). This corrects the variances of the $\{\tilde{\mu}_n, \tilde{\lambda}_n\}$ without significant change to their means.[8]

However, this does not work for $p(\tilde{\boldsymbol{\pi}}|\widetilde{\boldsymbol{X}})$, because it would change the mean of the LDP component, which is already correct, as well as its variance. To cope with this, we reexpress Eq. (22a) in the equivalent but extendable representation

$$p(\tilde{\boldsymbol{\pi}}|\widetilde{\boldsymbol{X}}) = \mathcal{GD}(R\cdot\tilde{\boldsymbol{\pi}};\{\phi_n,\psi_n\}) \,, \tag{26}$$

which uses the generalized Dirichlet distribution (5, 6) over proportion $(N+1)$-vectors:[9]

$$\mathcal{GD}(\boldsymbol{\pi};\{\phi_n,\psi_n\}) = \prod_{n=1}^{N} \frac{\pi_n^{\phi_n-1}(1-\pi_1-\ldots-\pi_n)^{\gamma_n}}{B(\phi_n,\psi_n)} \tag{27}$$

$$\gamma_n = \begin{cases} \psi_n - \phi_{n+1} - \psi_{n+1} & (1 \leq n \leq N-1) \\ \psi_N - 1 & (n = N) \end{cases}$$

$$\pi_{N+1} = 1 - \sum_{n=1}^{N} \pi_n \,.$$

where $B(\cdot,\cdot)$ is the beta function and there are $2N$ parameters $\{\phi_n, \psi_n : 1 \leq n < N\}$. This reduces to a Dirichlet distribution when $\gamma_n = 0$ $(1 \leq n < N)$. Eq. (26) includes the operator R, which reverses the order of indices:

$$(R\cdot\boldsymbol{\pi})_n = \pi_{N+1-n} \,. \tag{28}$$

---

[8] The means are very slightly altered because of a change in the weighting between the prior and posterior hyperparameters in Eqs. (23). This is as it should be, but has almost no effect because the hyperparameters are small and $\hat{Z}_N \gg 1$.

[9] We keep here the conventional notation in which the generalized Dirichlet is defined over components $1 \leq n \leq N+1$ rather than $0 \leq n \leq N$. The needed adjustment is made by Eq. (28).

Figure 3.11. continued

This reordering is needed because, in contrast with the Dirichlet distribution [Eq. (18)] which treats all components symmetrically, the generalized Dirichlet distribution is not invariant under permutation: Proportion variables located early in the ordering are independent of the ratios between variables located later in the ordering ("neutral"), but the converse does not hold (5). Given that this asymmetry is unavoidable and considering the special role of the LDP, it is most appropriate for it to be neutral with respect to the other variables, and R adjusts the ordering to this end.

We keep the posterior generalized Dirichlet means $\{\bar{\tilde{\pi}}_n\}$ and the $N-1$ variances $\{\widetilde{\Sigma}_{nn} : 1 \le n < N\}$ [Eqs. (19)] unchanged while doubling $\widetilde{\Sigma}_{NN}$ by choosing parameters $\{\phi_n, \psi_n\}$ such that

$$\bar{\tilde{\pi}}_n(\{\phi_n,\psi_n\}) = \frac{\alpha_n}{\alpha_\Sigma} \qquad (0 \le n \le N) \quad (29a)$$

$$\widetilde{\Sigma}_{nn}(\{\phi_n,\psi_n\}) = (1+\delta_{nN})\,\frac{\alpha_n(\alpha_\Sigma-\alpha_n)}{(\alpha_\Sigma)^2\,(\alpha_\Sigma+1)} \qquad (1 \le n \le N)\,, \quad (29b)$$

where the $\{\alpha_n\}$ are those specified by Eq. (23a). These constraints fix the $\phi_n(\{\alpha_n\})$ and $\psi_n(\{\alpha_n\})$ as described below. We use Eq. (26) with these values to compute the posterior distribution over $\tilde{\boldsymbol{\pi}}$.[10]

b. Extended tail of the $n = 0$ peak: Whether or not the analyzed data set was reflected, the extended right tail of the $n = 0$ component, comprising fraction $f_>$ of F^m_{res}, may have been pruned as described in Sec. I C 1. To compensate for this, the posterior distribution over $\tilde{\boldsymbol{\pi}}$ must be modified in accord with the portion that was analyzed. To do this we make the substitution $\tilde{\boldsymbol{\pi}} \to (\tilde{\boldsymbol{\pi}} - f_> \tilde{\boldsymbol{\epsilon}}_0)/(1-f_>)$, where $\tilde{\boldsymbol{\epsilon}}_n$ is the unit $(N+1)$-vector with n^{th} component equal one and all other components zero. (If there was no pruning, $f_> = 0$ and the substitution has no effect.)

Combining the adjustments gives

[10] For simplicity, we use Eq. (29b) even though it treats π_0 (in addition to π_N) asymmetrically relative to the $\{\pi_n : 1 \le n < N\}$. A slightly better approximation would be to replace this set of constraints with $\widetilde{\Sigma}_{nn} = c\,(1+\delta_{nN})\,\alpha_n(\alpha_\Sigma-\alpha_n)/[(\alpha_\Sigma)^2\,(\alpha_\Sigma+1)]$ $(0 \le n \le N)$, where c is a constant; this would treat all the $\{\pi_n : n \le 0 < N\}$ symmetrically. However, this would be more complicated and Eq. (29b) is adequate for our purposes.

Figure 3.11. continued

$$\left.\begin{aligned} p(\bar{\boldsymbol{\pi}}|\mathcal{X}^m_{\text{res}}) &= \mathcal{D}[(\bar{\boldsymbol{\pi}} - f_> \tilde{\boldsymbol{\epsilon}}_0)/(1-f_>); \{\alpha_n\}] \\ p(\tilde{\mu}_n|\tilde{\lambda}_n, \mathcal{X}^m_{\text{res}}) &= \mathcal{N}_>[\tilde{\mu}_n; \bar{\tilde{\mu}}_n, 1/(\beta_n \tilde{\lambda}_n)] \\ p(\tilde{\lambda}_n|\mathcal{X}^m_{\text{res}}) &= \mathcal{G}(\tilde{\lambda}_n; a_n, b_n) \end{aligned}\right\} \text{(fully resolved LDP)}$$

$$\left.\begin{aligned} p(\bar{\boldsymbol{\pi}}|\mathcal{X}^m_{\text{res}}) &= \mathcal{GD}[R\cdot(\bar{\boldsymbol{\pi}} - f_> \tilde{\boldsymbol{\epsilon}}_0)/(1-f_>); \{\phi_n(\{\alpha_n\}), \psi_n(\{\alpha_n\})\}] \\ p(\tilde{\mu}_n|\tilde{\lambda}_n, \mathcal{X}^m_{\text{res}}) &= \mathcal{N}_>[\tilde{\mu}^S_n; \bar{\tilde{\mu}}_n{}^S, 1/(\beta^S_n \tilde{\lambda}_n)] \\ p(\tilde{\lambda}_n|\mathcal{X}^m_{\text{res}}) &= \mathcal{G}(\tilde{\lambda}_n; a^S_n, b^S_n) \end{aligned}\right\} \text{(semi-resolved LDP)} \quad .$$

The mean of $\bar{\boldsymbol{\pi}}^m$ is

$$\bar{\bar{\pi}}_n = (1-f_>)\frac{\alpha_n}{\alpha_\Sigma} + f_> \delta_{n0} \tag{30}$$

and its covariance matrix is

$$\widetilde{\Sigma}_{nn'} = \frac{\alpha_n(\delta_{nn'}\alpha_\Sigma - \alpha_{n'})}{(\alpha_\Sigma)^2(\alpha_\Sigma+1)}(1-f_>)^2$$
$$\text{(fully resolved LDP)} \tag{31a}$$

$$\widetilde{\Sigma}_{R(n)R(n')} = \Sigma^{\mathcal{GD}}_{nn'}[\{\phi_n(\{\alpha_n\}), \psi_n(\{\alpha_n\})]\,(1-f_>)^2$$
$$\text{(semi-resolved LDP)}\,, \tag{31b}$$

where $\Sigma^{\mathcal{GD}}_{nn'}$ is defined in Eq. (32b) and the $\{\phi_n(\{\alpha_n\}), \psi_n(\{\alpha_n\})\}$ are computed as described below.

c. Parameterizing the Generalized Dirichlet Distribution by its Means and Variances: The means and covariance matrix of the generalized Dirichlet distribution over the proportion $(N+1)$-vector $(\pi_1, \ldots, \pi_{N+1})$, which is parameterized by the $2N$ parameters $\{\phi_n, \psi_n\}$ are (5)

$$\bar{\pi}_n = \frac{\phi_n}{\phi_n + \psi_n}\Psi_{n-1} \qquad (1 \le n \le N) \tag{32a}$$

$$\Sigma^{\mathcal{GD}}_{nn'} = \left(\frac{\phi_n + \delta_{nn'}}{\phi_n + \psi_n + 1}\Psi^1_{n-1} - \bar{\pi}_n\right)\bar{\pi}_{n'}$$
$$(1 \le n, n' \le N)\,, \tag{32b}$$

where

$$\Psi_n = \prod_{n'=1}^{n} \frac{\psi_{n'}}{\phi_{n'} + \psi_{n'}}$$
$$\Psi^1_n = \prod_{n'=1}^{n} \frac{\psi_{n'}+1}{\phi_{n'} + \psi_{n'} + 1}\,,$$

and $\bar{\pi}_{N+1}$ and $\Sigma^{\mathcal{GD}}_{n(N+1)} = \Sigma^{\mathcal{GD}}_{(N+1)n}$ $(1 \le n \le N+1)$ are determined by Eqs. (3) and (4). We reparameterize in terms of the $\{\bar{\pi}_n, \Sigma_{nn}\}$ by inverting Eqs. (32):

$$\phi_n = \frac{(\bar{\pi}_n^2 - \Psi^1_{n-1}\,\bar{\pi}_n)\,\bar{\pi}_n}{(\bar{\pi}_n)^2\,\Psi^1_{n-1} - \bar{\pi}_n^2\,\Psi_{n-1}} \tag{33a}$$

$$\psi_n = \frac{(\bar{\pi}_n\,\Psi^1_{n-1} - \bar{\pi}_n^2)\,(\bar{\pi}_n - \Psi_{n-1})}{(\bar{\pi}_n)^2\,\Psi^1_{n-1} - \bar{\pi}_n^2\,\Psi_{n-1}}\,, \tag{33b}$$

where

$$\bar{\pi}_n^2 = (\bar{\pi}_n)^2 + \Sigma_{nn}\,.$$

We determine the $\{\phi_n(\{\alpha_n\}, \psi_n(\{\alpha_n\})$ using Eqs. (33) with

$$\boldsymbol{\pi} \to R\cdot\bar{\bar{\boldsymbol{\pi}}}$$
$$\Sigma_{nn} \to \widetilde{\Sigma}_{R(n)R(n)}\,,$$

where $R(n) = N+1-n$, and the right-hand sides of the equations are from Eqs. (29).

Figure 3.11. continued

D. F^m_{amb}

To simplify notation, in this section we suppress m superscripts.

When the LDP is semi-resolved we must subtract $F_N/2$, the fraction coming from the left half of the LDP, from F_{lo} to get F_{amb}. Although $F_N/2$ is formally a stochastic function of $\bar{\pi}_N$, with little loss of accuracy we simplify the analysis by using $\mathrm{Mo}(\bar{\pi}_N)$, the maximum likelihood estimate of $\bar{\pi}_N$, to specify it deterministically. Recalling that F_{lo}, F_{res}, F_N, and F_{amb} are fractions relative to the total number of data points and that $\bar{\pi}_N$ is a proportion relative to $F_{\mathsf{res}} + F_N/2$, we have $F_N = (F_{\mathsf{res}} + F_N/2)\,\bar{\pi}_N$, so

$$\begin{aligned} F_{\mathsf{amb}} &= F_{\mathsf{lo}} && \text{(fully resolved LDP)} \\ \left.\begin{aligned} F_N &= \tfrac{\mathrm{Mo}(\bar{\pi}_N)}{1-\mathrm{Mo}(\bar{\pi}_N)/2} F_{\mathsf{res}} \\ F_{\mathsf{amb}} &= F_{\mathsf{lo}} - F_N/2 \end{aligned}\right\} & && \text{(semi-resolved LDP)}\,. \end{aligned}$$

F_{amb} contains contributions only from the unlabeled and highly replicated cells.

E. $p(F^m_{\mathsf{amb}}|\pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}})$

To simplify notation, in this section we suppress m superscripts.

Because π_{N+1} is normalized relative to the labeled cell fraction,

$$F_{\mathsf{amb}} = f_u + \pi_{N+1}\,(1 - f_u)\,. \tag{34}$$

f_u is a sufficient statistic with respect to $\mathcal{X}^{\mathcal{U}}$ and f_u and π_{N+1} are independent, so

$$\begin{aligned} p(F_{\mathsf{amb}}|\pi_{N+1}, \mathcal{X}^{\mathcal{U}}) &= \int p(F_{\mathsf{amb}}|\pi_{N+1}, f_u)\, p(f_u|\mathcal{X}^{\mathcal{U}})\, df_u \\ &= \int \delta\left[F_{\mathsf{amb}} - f_u - (1 - f_u)\,\pi_{N+1}\right]\, p(f_u|\mathcal{X}^{\mathcal{U}})\, df_u \\ &= p(f_u|\mathcal{X}^{\mathcal{U}})\Big|_{(F_{\mathsf{amb}}-\pi_{N+1})/(1-\pi_{N+1})}\,. \end{aligned} \tag{35}$$

The distribution $p(f_u|\mathcal{X}^{\mathcal{U}})$ was computed using background data from WT and p21 KO "unchased" mice that were sacrificed without administration of doxycycline (one WT at PD24 and one KO at PD25), or "short-chased" mice for which the chase time was insufficient for the generation of data points in the ambiguous region (PD24–27, four WT and one KO mouse; PD36–42, three WT and four KO mice). The unchased mice had a single, well-resolved $n = 0$ peak, and its data points could be easily distinguished from the low-fluorescence data points of the unlabeled cells. In the short-chased data, the LDP was fully resolved and there were no highly replicated cells, which again permitted unambiguous separation and counting of the labeled and unlabeled data points. Thus, each data set provided a value of f_u. One-way ANOVA and Kruskal-Wallis tests indicated that the probability of labeling is independent of genotype (determined by comparing WT and p21 KO mice separately for both the PD24–27 and PD36–42 chases) and of hair cycle stage (determined by comparing the PD24–27 and PD36–42 chases independently for both the WT and p21 KO mice). In all cases the probability for the null hypothesis of no difference was $p \geq 0.27$. Therefore, we pooled all 14 f_u values and used them to model $p(f_u|\mathcal{X}^{\mathcal{U}})$ as a beta distribution:

$$p(f_u|\mathcal{X}^{\mathcal{U}}) = \mathcal{B}e(f_u; \zeta_1^0, \zeta_2^0)\,, \tag{36}$$

where

$$\begin{aligned} &\mathcal{B}e(f; \zeta_1, \zeta_2) = \\ &\quad [\theta(f) - \theta(1 - f)]\,\frac{\Gamma(\zeta_1 + \zeta_2)}{\Gamma(\zeta_1)\,\Gamma(\zeta_1)}\, f^{\zeta_1 - 1}\,(1 - f)^{\zeta_2 - 1} \end{aligned} \tag{37}$$

Figure 3.11. continued

and $\theta(\cdot)$ is the Heaviside step function. It has mean and variance

$$\bar{f} = \frac{\zeta_1}{\zeta_1 + \zeta_2}$$
$$\sigma^2 = \frac{\zeta_1 \zeta_2}{(\zeta_1 + \zeta_2)^2(\zeta_1 + \zeta_2 + 1)} .$$

The maximum likelihood estimates were $\zeta_1^0 = 3.7$ and $\zeta_2^0 = 34.5$, corresponding to an average of $\sim 10\%$ unlabeled cells.

F. Combining the mouse-specific distributions

As per Eq. (10), we must evaluate

$$p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \propto p(\Sigma^{\mathcal{E}}) \prod_{m=1}^{M} \int p(\mathcal{X}^m | \tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}})\, p(\tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1} | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}})\, d\tilde{\boldsymbol{\pi}}^m\, d\pi^m_{N^m+1} \tag{38a}$$

$$\text{where} \quad p(\mathcal{X}^m | \tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}}) \propto p(F^m_{\mathsf{amb}} | \pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}}) \frac{p(\tilde{\boldsymbol{\pi}}^m | \mathcal{X}^m_{\mathsf{res}})}{p(\tilde{\boldsymbol{\pi}}^m)} . \tag{38b}$$

This is too complicated to compute exactly. To simplify we change variables back to $\boldsymbol{\pi}^m(\tilde{\boldsymbol{\pi}}^m, \pi^m_{N^m+1})$ so that Eq. (38a) becomes[11]

$$p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \propto \tag{39}$$
$$p(\Sigma^{\mathcal{E}}) \prod_{m=1}^{M} \int p(\mathcal{X}^m | \boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})\, p(\boldsymbol{\pi}^m | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}})\, d\boldsymbol{\pi}^m .$$

We next approximate both factors in this integral using constrained multinormal distributions.

1. *Approximating* $p(\boldsymbol{\pi}^m | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}})$

We approximate $p(\boldsymbol{\pi}^m | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}})$ to have the form

$$p(\boldsymbol{\pi}^m | \bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}}) = \mathcal{N}_{\Theta^m}(\boldsymbol{\pi}^m; \bar{\boldsymbol{\pi}}^{\mathcal{E}}_m, \Sigma^{\mathcal{E}}_m) , \tag{40}$$

where $\mathcal{N}_P(\boldsymbol{\pi}; \bar{\boldsymbol{\pi}}, \Sigma)$ is a constrained multinormal distribution and

$$\bar{\boldsymbol{\pi}}^{\mathcal{E}}_m \equiv U^m \cdot \bar{\boldsymbol{\pi}}^{\mathcal{E}}$$
$$\Sigma^{\mathcal{E}}_m \equiv U^m \cdot \Sigma^{\mathcal{E}} \cdot (U^m)^T ,$$

and U^m is the $(N^m+2) \times (N^{\mathcal{E}}+2)$ combination operator that projects proportions from the $(N^{\mathcal{E}}+2)$-vector space of $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ into the (N^m+2)-vector space of $\boldsymbol{\pi}^m$:

$$U^m_{nn'} = \tag{41}$$
$$\begin{cases} \delta_{nn'} & (0 \le n, n' \le N^m + 1) \\ 1 & (n = N^m + 1;\ N^m + 2 \le n' \le N^{\mathcal{E}} + 1) \\ 0 & (0 \le n \le N^m;\ N^m + 2 \le n' \le N^{\mathcal{E}} + 1) . \end{cases} \tag{42}$$

---

[11] Eq. (39) includes no Jacobian because Eq. (38b) has densities over $\tilde{\pi}^m$ both in the numerator and denominator.

Figure 3.11. continued

U^m adds the components $\{\pi_n^{\mathcal{E}} : N^m + 2 \leq n \leq N^{\mathcal{E}} + 1\}$, which are not resolved in $\boldsymbol{\pi}^m$, to $\pi_{N^m+1}^{\mathcal{E}}$ to generate the $(N^m + 2)$-vector that is the mean of $p(\boldsymbol{\pi}^m|\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}})$ and projects out the corresponding covariance tensor while preserving constraints Eq. (3) and (4):

$$\mathbf{1} \cdot \bar{\boldsymbol{\pi}}_m^{\mathcal{E}} = 1$$
$$\mathbf{1} \cdot \Sigma_m^{\mathcal{E}} = \Sigma_m^{\mathcal{E}} \cdot \mathbf{1} = 0 \,.$$

$\mathcal{N}_{\Theta^m}(\boldsymbol{\pi}; \bar{\boldsymbol{\pi}}, \Sigma) = C(\bar{\boldsymbol{\pi}}, \Sigma)\, \Theta^m(\boldsymbol{\pi}) \mathcal{N}_P(\boldsymbol{\pi}; \bar{\boldsymbol{\pi}}, \Sigma)$ is defined using the subspace-constrained multinormal distribution

$$\mathcal{N}_P(\boldsymbol{\pi}; \bar{\boldsymbol{\pi}}, \Sigma) = \frac{1}{\sqrt{|2\pi\Sigma|_+}}\, e^{-\frac{1}{2}(\boldsymbol{\pi}-\bar{\boldsymbol{\pi}})\cdot\Sigma^+\cdot(\boldsymbol{\pi}-\bar{\boldsymbol{\pi}})} \,,$$

the constraint function

$$\Theta^m(\boldsymbol{\pi}) = \delta\left(1 - \sum_{n=0}^{N^m+1} \pi_n\right) \prod_{n=0}^{N^m+1} \theta(\pi_n) \,, \tag{43}$$

and the appropriate normalizing constant $C(\bar{\boldsymbol{\pi}}, \Sigma)$. Care is needed because Eq. (4) implies that Σ is degenerate; Σ^+ denotes its pseudo-inverse and $|\Sigma|_+$ its pseudo-determinant. This ensures that it satisfies

$$\int_{-\infty}^{\infty} \mathcal{N}_P(\boldsymbol{\pi}; \bar{\boldsymbol{\pi}}, \Sigma)\, d\boldsymbol{\pi} = 1 \,,$$

where the negative and positive infinite integration limits imply that $d\boldsymbol{\pi}$ ranges over both negative and positive values while maintaining $\boldsymbol{\pi} \cdot \mathbf{1} = 1$ (i.e., $\mathbf{1} \cdot d\boldsymbol{\pi} = 0$).

2. *Approximating* $p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})$

We approximate $p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})$ up to a constant by a constrained multinormal distribution parameterized by its mean $\hat{\boldsymbol{\pi}}^m$ and covariance matrix $\widehat{\Sigma}^m$:

$$p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}}) \underset{\sim}{\propto} \mathcal{N}_{\Theta^m}(\boldsymbol{\pi}^m; \hat{\boldsymbol{\pi}}^m, \widehat{\Sigma}^m) \,, \tag{44}$$

where

$$\hat{\boldsymbol{\pi}}^m = \frac{\int \boldsymbol{\pi}^m\, p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})\, d\boldsymbol{\pi}^m}{\int p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})\, d\boldsymbol{\pi}^m}$$
$$\widehat{\Sigma}^m = \frac{\int (\boldsymbol{\pi}^m - \hat{\boldsymbol{\pi}}^m) \otimes (\boldsymbol{\pi}^m - \hat{\boldsymbol{\pi}}^m)\, p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})\, d\boldsymbol{\pi}^m}{\int p(\mathcal{X}^m|\boldsymbol{\pi}^m, \mathcal{X}^{\mathcal{U}})\, d\boldsymbol{\pi}^m} \,.$$

We perform these integrations in the $(\tilde{\boldsymbol{\pi}}^m, \pi_{N^m+1}^m)$ representation using the relationships

$$\pi_{N^m+1}^m = \frac{F_{\mathsf{amb}}^m - f_u^m}{1 - f_u^m}$$
$$\pi_n^m = \tilde{\pi}_n^m\,(1 - \pi_{N^m+1}^m) \qquad (0 \leq n \leq N) \,,$$

where the first line follows from Eq. (34) and the second from the different normalizations of $\boldsymbol{\pi}^m$ and $\tilde{\boldsymbol{\pi}}^m$. Using Eq. (38b) and the constancy of $p(\tilde{\boldsymbol{\pi}}^m)$ [Eq. (24)], the means are

$$\begin{aligned}
\hat{\pi}_{N^m+1}^m &= \frac{\int_0^1 \pi_{N^m+1}^m\, p(F_{\mathsf{amb}}^m|\pi_{N^m+1}^m, \mathcal{X}^{\mathcal{U}})\, d\pi_{N^m+1}^m}{\int_0^1 p(F_{\mathsf{amb}}^m|\pi_{N^m+1}^m, \mathcal{X}^{\mathcal{U}})\, d\pi_{N^m+1}^m} \\
&= \frac{\int_0^{F_{\mathsf{amb}}^m} \left(\frac{F_{\mathsf{amb}}^m - f_u^m}{1 - f_u^m}\right) \frac{\mathcal{B}e(f_u^m; \zeta_1^0, \zeta_2^0)}{(1-f_u^m)^2}\, df_u^m}{\int_0^{F_{\mathsf{amb}}^m} \frac{\mathcal{B}e(f_u^m; \zeta_1^0, \zeta_2^0)}{(1-f_u^m)^2}\, df_u^m} \\
&= \frac{F_{\mathsf{amb}}^m B(F_{\mathsf{amb}}^m; \zeta_1^0, \zeta_2^0 - 3) - B(F_{\mathsf{amb}}^m; \zeta_1^0 + 1, \zeta_2^0 - 3)}{B(F_{\mathsf{amb}}^m; \zeta_1^0, \zeta_2^0 - 2)}
\end{aligned} \tag{45}$$

Figure 3.11. continued

$$\hat{\pi}^m_n = \frac{\int_0^1 (1-\pi^m_{N^m+1})\, p(F^m_{\mathsf{amb}}|\pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}})\, d\pi^m_{N^m+1} \int \bar{\pi}^m_n\, p(\bar{\boldsymbol{\pi}}^m|\mathcal{X}^m_{\mathsf{res}})\, d\bar{\boldsymbol{\pi}}^m}{\int_0^1 p(F^m_{\mathsf{amb}}|\pi^m_{N^m+1}, \mathcal{X}^{\mathcal{U}})\, d\pi^m_{N^m+1}}$$
$$= (1-\hat{\pi}^m_{N^m+1})\, \bar{\bar{\pi}}^m_n \qquad (0 \le n \le N^m)\,, \tag{46}$$

where $B(\cdot\,;\cdot,\cdot)$ is the incomplete beta function and $\bar{\bar{\pi}}^m_n$ in Eq. (46) is defined in Eq. (30). To evaluate $\hat{\pi}^m_{N^m+1}$ we used Eqs. (35) with a Jacobian determined by Eq. (34) to change the integral over $\pi^m_{N^m+1}$ into one over f^m_u, and then applied Eqs. (36) and (37) and the fact that $\int_0^f \mathcal{B}e(f';\zeta_1,\zeta_2)\, df' = B(f;\zeta_1,\zeta_2)/B(\zeta_1,\zeta_2)$.

A similar calculation for the covariance matrix gives

$$\begin{aligned}
\widehat{\Sigma}^m_{(N^m+1)(N^m+1)} &= \tfrac{(F^m_{\mathsf{amb}})^2 I(F^m_{\mathsf{amb}};\zeta^0_1,\zeta^0_2-4) - 2F^m_{\mathsf{amb}} I(F^m_{\mathsf{amb}};\zeta^0_1+1,\zeta^0_2-4) + I(F^m_{\mathsf{amb}};\zeta^0_1+2,\zeta^0_2-4)}{B(F^m_{\mathsf{amb}};\zeta^0_1,\zeta^0_2-2)} - (\hat{\pi}^m_{N^m+1})^2 \\
\widehat{\Sigma}^m_{n(N^m+1)} &= -\bar{\bar{\pi}}^m_n\, \widehat{\Sigma}^m_{(N^m+1)(N^m+1)} \qquad (0 \le n \le N^m) \\
\widehat{\Sigma}^m_{nn'} &= \widetilde{\Sigma}^m_{nn'}\, \widehat{\Sigma}^m_{(N^m+1)(N^m+1)} + \widetilde{\Sigma}^m_{nn'}(1-\hat{\pi}^m_{N^m+1})^2 + \widehat{\Sigma}^m_{(N^m+1)(N^m+1)}\, \bar{\bar{\pi}}^m_n\, \bar{\bar{\pi}}^m_{n'} \qquad (0 \le n, n' \le N^m)\,,
\end{aligned} \tag{47}$$

where $\widetilde{\Sigma}^m$ is given by Eqs. (31). As must be so, $\mathbf{1}\cdot\widehat{\Sigma} = \widehat{\Sigma}\cdot\mathbf{1} = 0$.

3. *Convolving the distributions*

Combining Eqs. (39), (40), and (44) gives

$$p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \underset{\sim}{\propto} p(\Sigma^{\mathcal{E}}) \times$$
$$\prod_{m=1}^{M} \int \mathcal{N}_{\Theta^m}(\boldsymbol{\pi}^m; \bar{\boldsymbol{\pi}}^{\mathcal{E}}_m, \Sigma^{\mathcal{E}}_m)\, \mathcal{N}_{\Theta^m}(\boldsymbol{\pi}^m; \hat{\boldsymbol{\pi}}^m, \widehat{\Sigma}^m)\, d\boldsymbol{\pi}^m\,.$$

The integral is complicated by the inequality constraints. However, these are only needed when the the integrand has significant support in the negative region, which only occurs if $\bar{\boldsymbol{\pi}}^{\mathcal{E}}_n$ is of the same magnitude as its standard deviation. Thus, with little loss of accuracy we approximate the integral by ignoring the inequality constraints while maintaining the proportion summation constraints. These give

$$p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \underset{\sim}{\propto} p(\Sigma^{\mathcal{E}}) \times$$
$$\prod_{m=1}^{M} \int_{-\infty}^{\infty} \mathcal{N}_P(\boldsymbol{\pi}^m; \bar{\boldsymbol{\pi}}^{\mathcal{E}}_m, \Sigma^{\mathcal{E}}_m)\, \mathcal{N}_P(\boldsymbol{\pi}^m; \hat{\boldsymbol{\pi}}^m, \widehat{\Sigma}^m)\, d\boldsymbol{\pi}^m$$
$$= p(\Sigma^{\mathcal{E}}) \prod_{m=1}^{M} \mathcal{N}_P(\bar{\boldsymbol{\pi}}^{\mathcal{E}}_m, \hat{\boldsymbol{\pi}}^m, \widehat{\Sigma}^m + \Sigma^{\mathcal{E}}_m)\,. \tag{48}$$

The last line follows because each integral is the convolution of two subspace-constrained multinormal distributions over the constrained subspace.[12]

---

[12] Recall that the standard result is $\int_{-\infty}^{\infty} \mathcal{N}(x;\bar{x}_a,\Sigma_a)\, \mathcal{N}(x;\bar{x}_b,\Sigma_b)\, dx = \mathcal{N}(\bar{x}_a;\bar{x}_b,\Sigma_a+\Sigma_b)$. The analogous result follows for integration over the subspace.

Figure 3.11. continued

4. Combining the distributions

Taking the product of the subspace-constrained multinormal distributions in Eq. (48) gives

$$\begin{aligned} p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \boldsymbol{\mathcal{X}}^{\mathcal{E}}, \boldsymbol{\mathcal{X}}^{\mathcal{U}}) &\underset{\sim}{\propto} \\ & p(\Sigma^{\mathcal{E}}) \prod_{m=1}^{M} \left(|2\pi(\widehat{\Sigma}^m + \Sigma_m^{\mathcal{E}})|_+ \right)^{-1/2} \times \\ & e^{-\frac{1}{2}\sum_{m=1}^{M} (\boldsymbol{\pi}_m^{\mathcal{E}} - \hat{\boldsymbol{\pi}}^m) \cdot (\widehat{\Sigma}^m + \Sigma_m^{\mathcal{E}})^+ \cdot (\boldsymbol{\pi}_m^{\mathcal{E}} - \hat{\boldsymbol{\pi}}^m)} \\ = p(\Sigma^{\mathcal{E}})\, & f(\Sigma^{\mathcal{E}}, \{\hat{\boldsymbol{\pi}}^m, \widehat{\Sigma}^m\}) \times \\ & e^{-\frac{1}{2}(\boldsymbol{\pi}^{\mathcal{E}} - \boldsymbol{\pi}_{\Sigma}^{\mathcal{E}}) \cdot (\Sigma_{\Sigma}^{\mathcal{E}})^+ \cdot (\boldsymbol{\pi}^{\mathcal{E}} - \boldsymbol{\pi}_{\Sigma}^{\mathcal{E}})} , \end{aligned} \tag{49}$$

where

$$(\Sigma_{\Sigma}^{\mathcal{E}})^+ = P \cdot \left[\sum_{m=1}^{M} (U^m)^T \cdot (\widehat{\Sigma}^m + \Sigma_m^{\mathcal{E}})^+ \cdot U^m \right] \cdot P, \tag{50a}$$

$$\begin{aligned} \boldsymbol{\pi}_{\Sigma}^{\mathcal{E}} &= \mathbf{1}/(\mathbf{1}\cdot\mathbf{1}) + \\ & \Sigma_{\Sigma}^{\mathcal{E}} \cdot \sum_{m=1}^{M} (U^m)^T \cdot (\widehat{\Sigma}^m + \Sigma_m^{\mathcal{E}})^+ \cdot (\hat{\boldsymbol{\pi}}^m + \Delta\boldsymbol{\pi}_0^m) , \end{aligned} \tag{50b}$$

$$\begin{aligned} P &= I - \mathbf{1} \otimes \mathbf{1}/(\mathbf{1}\cdot\mathbf{1}) , \\ \Delta\boldsymbol{\pi}_0^m &= \mathbf{1}^m/(N^m+2) - U^m \cdot \mathbf{1}^{\mathcal{E}}/(N^{\mathcal{E}}+2) , \end{aligned}$$

and $\mathbf{1}^{\mathcal{E}}$ and $\mathbf{1}^m$ are the $(N^{\mathcal{E}}+2)$- and (N^m+2)-vectors, respectively, having all elements equal to one.[13] $f(\Sigma^{\mathcal{E}}, \{\hat{\boldsymbol{\pi}}^m, \widehat{\Sigma}^m\})$ absorbs the prefactor and the $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$-independent term that remains in the exponent after completing the square in $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$. U^m and the projection operator P in Eq. (50a) embed the $(N^m+2)\times(N^m+2)$ covariance matrix $\Sigma_{\Sigma}^{\mathcal{E}}$ into the $(N^{\mathcal{E}}+2)\times(N^{\mathcal{E}}+2)$ tensor space.

G. Empirical Bayes treatment of $\Sigma^{\mathcal{E}}$

A fully Bayesian treatment of the dependence of $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}, \Sigma^{\mathcal{E}} | \boldsymbol{\mathcal{X}}^{\mathcal{E}}, \boldsymbol{\mathcal{X}}^{\mathcal{U}})$ on $\Sigma^{\mathcal{E}}$ is meaningless in the absence of a realistic prior distribution because there is not enough data to form a good estimate of the distribution over this variate. Therefore, we use an empirical Bayesian prior $p(\Sigma^{\mathcal{E}})$ that, within restraints, maximizes the likelihood of the data. So as not to overparameterize the estimate relative to the amount of available data, we restrict $\Sigma^{\mathcal{E}}$ to have same form as the $\widehat{\Sigma}^m$ matrices, each of which is determined by the (N^m+2)-vector $\hat{\boldsymbol{\pi}}^m(N^m+1)$, $\widehat{\Sigma}^m_{(N^m+1)(N^m+1)}$, the variance of the highly replicated component, and the $(N^m+1)\times(N^m+1)$ $\widetilde{\Sigma}^m$ matrix, which itself is determined by its diagonal elements. This form respects the anticorrelation between variations of $\bar{\pi}^{\mathcal{E}}_{N^{\mathcal{E}}+1}$ and the sum of the division-resolved components, which is reflected in Eqs. (47). This will be important below when we analyze the best-fit of the two-population Poisson model to the

---

[13] To pass from the second to the third line of Eq. (49) we expand $\bar{\pi}_m^{\mathcal{E}} = U^m \cdot [\mathbf{1}^{\mathcal{E}}/(N^{\mathcal{E}}+2) + \delta\bar{\pi}^{\mathcal{E}}]$ and $\hat{\pi}^m = \mathbf{1}^m/(N^m+2) + \delta\hat{\pi}^m$, which gives $(\bar{\pi}_m^{\mathcal{E}} - \hat{\pi}^m) = (U^m \cdot \delta\bar{\pi}^{\mathcal{E}} - \delta\hat{\pi}^m - \Delta\boldsymbol{\pi}_0^m)$. We then complete the square in the exponent to bring it to the form $-(1/2)\,(\delta\bar{\pi}^{\mathcal{E}} - \delta\bar{\pi}_{\Sigma}^{\mathcal{E}}) \cdot (\Sigma_{\Sigma}^{\mathcal{E}})^+ \cdot (\delta\bar{\pi}^{\mathcal{E}} - \delta\bar{\pi}_{\Sigma}^{\mathcal{E}})$ and use $\bar{\pi}_{\Sigma}^{\mathcal{E}} = \mathbf{1}^{\mathcal{E}}/(N^{\mathcal{E}}+2) + \delta\bar{\pi}_{\Sigma}^{\mathcal{E}}$ to get Eq. (49). The advantage of this approach is that $\mathbf{1}^{\mathcal{E}} \cdot \delta\bar{\pi}^{\mathcal{E}} = \mathbf{1}^{\mathcal{E}} \cdot \delta\bar{\pi}_{\Sigma}^{\mathcal{E}} = \mathbf{1}^m \cdot U^m \cdot \delta\bar{\pi}^{\mathcal{E}} = \mathbf{1}^{N^m} \cdot \delta\hat{\pi}^m = 0$, so all the "$\delta$-vectors" are in the null-space of covariance matrices satisfying the constraint Eq. (4); this simplifies expressing the solution in terms of parameters $\bar{\boldsymbol{\pi}}_{\Sigma}^{\mathcal{E}}$ and $\Sigma_{\Sigma}^{\mathcal{E}}$ that satisfy the proportion vector and covariance matrix constraints.

Figure 3.11. continued

data. [If the N^m were all equal, the maximum likelihood estimator would have precisely this form, but the inclusion of the U^m in Eqs. (50) makes it an approximation.]

That is, we consider $\Sigma^{\mathcal{E}}$ of the form[14]

$$\Sigma^{\mathcal{E}}_{(N^{\mathcal{E}}+1)(N^{\mathcal{E}}+1)} = \sigma^2_{N^{\mathcal{E}}+1} \tag{51a}$$

$$\Sigma^{\mathcal{E}}_{n(N^{\mathcal{E}}+1)} = -(\tilde{\bar{\pi}}^{\mathcal{E}}_{\Sigma})_n \, \sigma^2_{N^{\mathcal{E}}+1} \quad (0 \le n \le N^{\mathcal{E}}) \tag{51b}$$

$$\Sigma^{\mathcal{E}}_{nn'} = \tilde{\Sigma}^{\mathcal{E}}_{nn'} + \sigma^2_{N^{\mathcal{E}}+1} \, (\tilde{\bar{\pi}}^{\mathcal{E}}_{\Sigma})_n \, (\tilde{\bar{\pi}}^{\mathcal{E}}_{\Sigma})_{n'} \quad (0 \le n, n' \le N^{\mathcal{E}}) \tag{51c}$$

where $\sigma^2_{N^{\mathcal{E}}+1}$ is an independent parameter and the proportion $(N^{\mathcal{E}}+1)$-vector and $(N^{\mathcal{E}}+1) \times (N^{\mathcal{E}}+1)$ covariance matrix within the division-resolved subspace are

$$(\tilde{\bar{\pi}}^{\mathcal{E}}_{\Sigma})_n = \frac{(\bar{\pi}^{\mathcal{E}}_{\Sigma})_n}{1 - (\bar{\pi}^{\mathcal{E}}_{\Sigma})_{N^{\mathcal{E}}+1}} \qquad (0 \le n \le N^{\mathcal{E}})$$

$$\tilde{\Sigma}^{\mathcal{E}} = P \cdot \mathsf{diag}(\sigma^2_0, \ldots, \sigma^2_{N^{\mathcal{E}}}) \cdot P \, . \tag{52}$$

The structure of Eqs. (51) and the projection operator in Eq. (52) ensure that $\Sigma^{\mathcal{E}}$ and $\tilde{\Sigma}^{\mathcal{E}}$ satisfy Eq. (4) within their respective vector spaces.

Since $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}})$ is non-informative, maximizing the likelihood is equivalent to maximizing Eq. (49) with the $p(\Sigma^{\mathcal{E}})$ term omitted over the independent parameters $\{\sigma^2_0, \ldots, \sigma^2_{N^{\mathcal{E}}}, \Sigma^{\mathcal{E}}_{(N^{\mathcal{E}}+1)(N^{\mathcal{E}}+1)}\}$.[15] We designate the maximizer as $\Sigma^{\mathcal{E}\star}_{\Sigma}$. Combining this with Eq. (49) we get

$$p(\bar{\boldsymbol{\pi}}^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) \approx \mathcal{N}(\bar{\boldsymbol{\pi}}^{\mathcal{E}}; \bar{\boldsymbol{\pi}}^{\mathcal{E}\star}_{\Sigma}, \Sigma^{\mathcal{E}\star}_{\Sigma}) \, , \tag{53}$$

where $\bar{\boldsymbol{\pi}}^{\mathcal{E}\star}_{\Sigma}$ and $\Sigma^{\mathcal{E}\star}_{\Sigma}$ are determined by Eqs. (50) with $\Sigma^{\mathcal{E}} \to \Sigma^{\mathcal{E}\star}$.

H. Summary

In summary the overall procedure consists of: 1) identifying the live bulge cells and collecting the mouse datasets $\{\mathcal{X}^m\}$ and the background subtraction data set $\mathcal{X}^{\mathcal{U}}$, 2) dissecting each $\mathcal{X}^m$ into $\mathcal{X}^m_{\mathsf{lo}}$ and $\mathcal{X}^m_{\mathsf{res}}$, 3) iteratively determining the mean-field hyperparameters, 4) computing the maximum likelihood estimate of the unlabeled cell distribution hyperparameters, 5) determining the $\{\bar{\tilde{\boldsymbol{\pi}}}^m\}$ and $\{\tilde{\Sigma}^m\}$, the means and covariance matrices of the $\{\tilde{\boldsymbol{\pi}}^m\}$, from the analytic properties of the Dirichlet or generalized Dirichlet distributions [Eqs. (30) and (31)], 6) analytically determining the $\{\hat{\boldsymbol{\pi}}^m\}$ and $\{\hat{\Sigma}^m\}$, the means and covariance matrices of the $\{\boldsymbol{\pi}^m\}$, from the $\{\bar{\tilde{\boldsymbol{\pi}}}^m\}$ and $\{\tilde{\Sigma}^m\}$ [Eqs. (45), (46), and (47)], 7) algebraically determining $\hat{\boldsymbol{\pi}}^{\mathcal{E}}_{\Sigma}$ and $\Sigma^{\mathcal{E}}_{\Sigma}$, the mean and variance of the conditional posterior over $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ and $\Sigma^{\mathcal{E}}$, from the $\{\hat{\boldsymbol{\pi}}^m\}$ and $\{\hat{\Sigma}^m\}$ [Eqs. (50)], and 8) fixing $\Sigma^{\mathcal{E}}$ at its maximum likelihood estimator $\Sigma^{\mathcal{E}\star}$ to determine $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}} | \mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$.

The bar charts in Figures 2 and S2 display the components of $\hat{\boldsymbol{\pi}}^{\mathcal{E}\star}_{\Sigma}$ and the projected standard deviations $\sqrt{(\Sigma^{\mathcal{E}\star}_{\Sigma})_{nn}}$.

II. MODELING PROLIFERATION DYNAMICS DURING PD21-47

To better understand how p21 KO perturbs bulge replication, we mathematically modeled cell proliferation across the entire hair cycle (i.e., from PD21-47) as a function of genotype. We first tested a homogeneous

---

[14] Eqs. (47) contain additional factors multiplying $\tilde{\Sigma}^m$, but these are not needed here since $\tilde{\Sigma}^{\mathcal{E}}$ is itself a variable matrix.

[15] We enforce both the non-negativity and summation constraints on $\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\Sigma}$ during the maximization.

Figure 3.11. continued

Poisson model in which the cells divide and are lost (e.g., by apoptosis or export) with (possibly time-varying) homogeneous replication and loss rates that are the same for all the cells. This model (Sec. II A) produced a good fit to the PD21–47 WT data (Figure 2F)—with a mean number of replications of $\kappa_r^{\mathsf{wt}} = 3.14$. In contrast (Figure 2G), the homogeneous model could not provide a good fit to the KO data, which displayed an excess proportion of highly replicated cells. Thus, while the analysis can not prove that the WT cell population is homogeneous with regard to division and loss rates, it does prove that the KO population is not.

To better characterize the excess of highly replicated KO cells, we tested whether the addition of a second homogeneous population with greater replication, which did not effect the coexisting WT population, could explain the data (Sec. II B). A very good fit to the data was achieved; it implied that by PD47 $\sim 40\%$ of the population was comprised of cells that had replicated at least three times more than the WT cells over the whole cycle.

A. Homogeneous Poisson Model

We define the number of cells at time t that have divided n times in the homogeneous Poisson model as ${}^{\mathsf{p}}\eta_n(t; k_r, k_l)$, which satisfies

$$\frac{d\,{}^{\mathsf{p}}\eta_n[t; k_r, k_l]}{dt} = \tag{54}$$
$$2k_r(t)\,{}^{\mathsf{p}}\eta_{n-1}(t; k_r, k_l) - [k_r(t) + k_l(t)]\,{}^{\mathsf{p}}\eta_n(t; k_r, k_l)$$

where $k_r(t)$ and $k_l(t)$ are the time-varying rates of division and loss, respectively. Specifying that there are η_0 undivided cells at $t = t_0$,

$${}^{\mathsf{p}}\eta_n[t_0; k_r, k_l] = \delta_{n0}\ \eta_0\ . \tag{55}$$

This has the solution

$$\begin{aligned} {}^{\mathsf{p}}\eta_n[t; k_r, k_l] &= \\ &\eta_0 \frac{\left[2\int_{t_0}^{t} k_r(t')dt'\right]^n}{n!} \exp\left\{-\int_{t_0}^{t} [k_r(t') + k_l(t')]\,dt'\right\} \\ &= \eta_0 \mathcal{P}(n; 2\kappa_r)\, e^{\Delta\kappa}\ , \end{aligned} \tag{56}$$

where

$$\mathcal{P}(n; \mu) = \frac{\mu^n}{n!} e^{-\mu}$$

is the Poisson distribution and

$$\begin{aligned} \kappa_r &\equiv \int_{t_0}^{t} k_r(t')\,dt' \\ \kappa_l &\equiv \int_{t_0}^{t} k_l(t')\,dt' \\ \Delta\kappa &\equiv \kappa_r - \kappa_l \end{aligned} \tag{57}$$

(We leave the t_0 and t dependence of κ_r, κ_l, and $\Delta\kappa$ implicit.)

The predicted proportion of cells that have replication number n at time t is

$$\begin{aligned} {}^{\mathsf{p}}\pi_n(\kappa_r) &= \frac{{}^{\mathsf{p}}\eta_n[t; k_r, k_l]}{{}^{\mathsf{p}}\eta_{\mathsf{tot}}[t; k_r, k_l]} \\ &= \mathcal{P}(n; 2\kappa_r)\ , \end{aligned} \tag{58}$$

Figure 3.11. continued

where

$$ {}^{x}\eta_{\text{tot}}(t;k_r,k_l) = \sum_n {}^{x}\eta_n[t;k_r,k_l] = \eta_0\, e^{\Delta\kappa} \tag{59} $$

is the total number of cells at time t.

${}^{x}\boldsymbol{\pi}(\kappa_r)$ has three notable features: 1) It does not depend on $k_l(t)$. 2) Instead of being a functional of $k_r(t)$, it is a function only of the time-integrated value κ_r. 3) Assuming, as above, that any loss is replication-independent, the mean number of replications that have occurred for the cells that are present at time t,

$$ {}^{x}\langle n\rangle(\kappa_r) = \sum_n n\ {}^{x}\pi_n(\kappa_r) = 2\kappa_r\ , \tag{60} $$

is *twice* κ_r.

The first point follows from the assumption that the rate of cell loss does not depend on its replication history: both highly- and rarely-replicated cells are lost at the same rate, and this does not alter their ratio.[16] The second point indicates that in this model the replication proportion data $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ can not discriminate the temporal order of any periods of high and low replication replication within the hair cycle. The third point emphasizes that we must interpret the replication data with care: ${}^{x}\langle n\rangle(\kappa_r)$ is a *retrospective* statistic that describes the average number of times that a cell existing at time t *has* divided since t_0. Compare it with the *prospective* statistic ${}^{x}\langle r\rangle(\kappa_r)$ that specifies the average number of times that a cell existing at t_0 and a linear chain of descendants *will* divide by time t (assuming that the chain is not broken by cell loss). Since the mean waiting time between cell divisions occurring with Poisson statistics is $(t-t_0)/\kappa_r$, this is

$$ {}^{x}\langle r\rangle(\kappa_r) = \kappa_r = \frac{{}^{x}\langle n\rangle(\kappa_r)}{2}\ . $$

1. *Statistical estimation of* ${}^{x}\kappa_r^{\text{wt}}$

We would like to use Eq. (58) with $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})$ to estimate the most probable range of κ_r. Two problems must be addressed: 1) ${}^{x}\boldsymbol{\pi}(\kappa_r)$ is an infinite-dimensional vector while $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ is a $(N^{\mathcal{E}}+2)$-vector, and 2) $p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})$ was computed using a completely uninformative prior, while the homogeneous Poisson model implicitly assumes that the proportion vector is limited to the "trajectory" of values that are consistent with Eq. (58).

To solve the first problem, we project the infinite-dimensional vector ${}^{x}\boldsymbol{\pi}(\kappa_r)$ into the $(N^{\mathcal{E}}+2)$-dimensional subspace by

$$ {}^{x}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r) \equiv U^{\mathcal{E}}\cdot {}^{x}\boldsymbol{\pi}(\kappa_r)\ , \tag{61} $$

where $U^{\mathcal{E}}$ is the infinite-dimensional generalization of Eq. (42):

$$ U^{\mathcal{E}}_{nn'} = \begin{cases} \delta_{nn'} & (0\le n,n'\le N^{\mathcal{E}}+1) \\ 1 & (n=N^{\mathcal{E}}+1;\ N^{\mathcal{E}}+2\le n'<\infty) \\ 0 & (0\le n\le N^{\mathcal{E}};\ N^{\mathcal{E}}+2\le n'<\infty)\ . \end{cases} $$

To solve the second, we introduce a prior over $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ that is uniform within the subspace of proportion vectors that are consistent with the Poisson model. This subspace is defined by the *solution trajectory* ${}^{x}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)$ given by Eq. (58) with $0\le\kappa_r<\infty$ and is a curvilinear path through $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$-space that begins at ${}^{x}\boldsymbol{\pi}^{\mathcal{E}}(0)=\boldsymbol{\epsilon}_0$ and ends at ${}^{x}\boldsymbol{\pi}^{\mathcal{E}}(\infty)=\boldsymbol{\epsilon}_{N^{\mathcal{E}}+1}$, where $\boldsymbol{\epsilon}_n$ is the unit $(N^{\mathcal{E}}+2)$-vector with n^{th} component equal one and all other

---

[16] This assumption could be violated if replication and export were mechanistically linked. However, since these processes occur primarily at different times in the hair cycle (see Sec. III), this is unlikely.

Figure 3.11. continued

components zero. (Note that κ_r varies because of changes in $k_r(t')$, not changes in t, which is fixed.) The trajectory can be expressed as

$$ {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r) = \boldsymbol{\mathcal{P}}^{\mathcal{E}}(\kappa_r) \cdot \boldsymbol{\epsilon}_0 \ , \tag{62} $$

where the propagation matrix $\boldsymbol{\mathcal{P}}^{\mathcal{E}}(\kappa_r)$ is

$$ \boldsymbol{\mathcal{P}}^{\mathcal{E}}(\kappa_r) = U^{\mathcal{E}} \cdot \boldsymbol{\mathcal{P}}(\kappa_r) \tag{63a} $$

$$ \mathcal{P}_{n,n'}(\kappa_r) = \begin{cases} \frac{(2\kappa_r)^{n-n'}}{(n-n')!}\, e^{-2\kappa_r} & (0 \le n' \le n < \infty) \\ 0 & (0 \le n < n' < \infty) \ . \end{cases} \tag{63b} $$

Because all the highly-replicated cells are compressed into the last component, $\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)/\partial\kappa_r \to 0$ as $\kappa_r \to \infty$,[17] so a density that is uniform over κ_r is highly nonuniform in $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$-space and is not appropriate.

a. The prior ${}^{\mathcal{P}}p(\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}})$*:* To define a uniform-density prior we need to map the region surrounding the solution trajectory into the $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ coordinates so densities can be defined. For this purpose we parameterize $\text{int}(\Pi^{\mathcal{E}})$, the interior of the constrained space $\Pi^{\mathcal{E}} \equiv \{\bar{\boldsymbol{\pi}}^{\mathcal{E}}\}$, in terms of κ_r and $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}} \in \partial\Pi_0^{\mathcal{E}}$, a proportion-vector on the portion of the constraint boundary (i.e., where at least one component of $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}}$ equals zero) that has a positive 0^{th} component, $(\bar{\boldsymbol{\pi}}_0^{\mathcal{E}})_0 > 0$. By using $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}}$ as an alternative initial condition in Eq. (62) we generate the mapping

$$ \begin{aligned} \bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}}) = & \boldsymbol{\mathcal{P}}^{\mathcal{E}}(\kappa_r) \cdot \bar{\boldsymbol{\pi}}_0^{\mathcal{E}} \\ & [\bar{\boldsymbol{\pi}}_0^{\mathcal{E}} \in \partial\Pi_0^{E} \, , \bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}}) \in \text{int}(\Pi^{\mathcal{E}})] \ . \end{aligned} $$

Each set of points $\bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}})$, with $0 < \kappa_r < \infty$ defines a trajectory, and

$$ \bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r, \boldsymbol{\epsilon}_0) = {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r) \tag{64} $$

defines the solution trajectory.

$\boldsymbol{\mathcal{P}}^{\mathcal{E}}(\kappa_r)$ is a lower-triangular matrix with non-zero diagonal elements so it is invertible with $(\boldsymbol{\mathcal{P}}^{\mathcal{E}})^{-1}(\kappa_r) = \boldsymbol{\mathcal{P}}^{\mathcal{E}}(-\kappa_r)$. Moreover, the mapping is invertible for any $\bar{\boldsymbol{\pi}}^{\mathcal{E}} \in \text{int}(\Pi^{\mathcal{E}})$ because there is a unique κ_r such that $(\boldsymbol{\mathcal{P}}^{\mathcal{E}})^{-1}(\kappa_r) \cdot \bar{\boldsymbol{\pi}}^{\mathcal{E}}$ lies on $\partial\Pi_0^{\mathcal{E}}$.[18] Therefore, the family of all trajectories covers $\text{int}(\Pi^{\mathcal{E}})$ and we have the global one-to-one mapping

$$ \bar{\boldsymbol{\pi}}^{\mathcal{E}} \leftrightarrow (\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}}) \qquad (\bar{\boldsymbol{\pi}}^{\mathcal{E}} \in \Pi^{E}) \ . \tag{65} $$

Using this parameterization, we restrict the posterior distribution to solutions satisfying Eq. (58) by using a prior of the form

$$ {}^{\mathcal{P}}p(\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}}) \propto \delta(\bar{\boldsymbol{\pi}}_0^{\mathcal{E}} - \boldsymbol{\epsilon}_0)\, F(\kappa_r) \ , \tag{66} $$

where $F(\kappa_r)$ is determined by the requirement that the integral of ${}^{\mathcal{P}}p(\kappa_r, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}})$ over an infinitesimal $(N^{\mathcal{E}}+1)$-dimensional orthotope (the generalization of a rectangle to higher dimensions) surrounding a portion of the trajectory in $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$-space be constant as the orthotope is moved along the trajectory while maintaining its orientation parallel to the trajectory tangent vector

$$ \boldsymbol{v}^{\mathcal{E}}(\kappa_r) = \frac{\partial\bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r, \boldsymbol{\epsilon}_0)/\partial\kappa_r}{\left|\partial\bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r, \boldsymbol{\epsilon}_0)/\partial\kappa_r\right|} \ . $$

It is convenient to introduce $s(\kappa_r)$, the path-length along the solution trajectory, which is computed by integrating

---

[17] Because ${}^{\mathcal{P}}\boldsymbol{\pi}(\kappa_r)$ is governed by a Poisson distribution with mean $2\kappa_r$ and standard deviation $\sqrt{2\kappa_r}$, ${}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r) \approx \epsilon_{N^{\mathcal{E}}+1}$ when $2\kappa_r - \sqrt{2\kappa_r} \gg N^{\mathcal{E}} + 1$.

Figure 3.11. continued

$$
\begin{aligned}
ds &= d\kappa_r \left| \frac{\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)}{\partial \kappa_r} \right| \\
&= d\kappa_r \sqrt{\left(\frac{\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)}{\partial \kappa_r} \right)^T \cdot \frac{\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)}{\partial \kappa_r}} \\
&= d\kappa_r\, 2\, e^{-2\kappa_r} \sqrt{\sum_{n=0}^{N^{\mathcal{E}}} \left[\frac{(n-2\kappa_r)(2\kappa_r)^{n-1}}{n!} \right]^2 + \left(\sum_{n=N^{\mathcal{E}}+1}^{\infty} \frac{(n-2\kappa_r)(2\kappa_r)^{n-1}}{n!} \right)^2} \\
&= d\kappa_r\, 2\, e^{-2\kappa_r} \sqrt{\sum_{n=0}^{N^{\mathcal{E}}} \left[\frac{(n-2\kappa_r)(2\kappa_r)^{n-1}}{n!} \right]^2 + \left[\frac{(2\kappa_r)^{N^{\mathcal{E}}}}{N^{\mathcal{E}}!} \right]^2}
\end{aligned}
\tag{67}
$$

with the boundary condition $s(0) = 0$. $s(\kappa_r)$ is invertible, so $\kappa_r(s)$ is uniquely defined on the solution trajectory. To define the orthotope surrounding trajectory point ${}^{\mathcal{P}}\boldsymbol{\pi}(\kappa_r, \boldsymbol{\epsilon}_0)$ we introduce $N^{\mathcal{E}}$ basis vectors $\{\boldsymbol{x}_{\perp,a}(\kappa_r) : 1 \le a \le N^{\mathcal{E}}\}$ that span the orthogonal complement to $\boldsymbol{v}^{\mathcal{E}}(\kappa_r)$ and their corresponding *local* coordinates $\{\Delta\pi^{\mathcal{E}}_{\perp,a} : 1 \le a \le N^{\mathcal{E}}\}$. Because $\boldsymbol{v}^{\mathcal{E}}(\kappa_r)$ changes direction along the trajectory, the $\{\boldsymbol{x}_{\perp,a}(\kappa_r)\}$ also rotate as κ_r varies.[19] Thus, in the vicinity of $\bar{\boldsymbol{\pi}}^{\mathcal{E}}[\kappa_r(s), \boldsymbol{\epsilon}_0]$ we have the locally invertible mapping

$$
\bar{\boldsymbol{\pi}}^{\mathcal{E}} \leftrightarrow (s, \Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp}) \qquad (\bar{\boldsymbol{\pi}}^{\mathcal{E}} \approx \bar{\boldsymbol{\pi}}^{\mathcal{E}}[\kappa_r(s), \boldsymbol{\epsilon}_0]) \tag{68}
$$

specified by

$$
\bar{\boldsymbol{\pi}}^{\mathcal{E}}(s, \Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp}) = \bar{\boldsymbol{\pi}}^{\mathcal{E}}[\kappa_r(s), \boldsymbol{\epsilon}_0] + \sum_{a=1}^{N^{\mathcal{E}}} \Delta\pi^{\mathcal{E}}_{\perp,a}\, \boldsymbol{x}_{\perp,a}(\kappa_r) \;,
$$

where $\Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp} \equiv \{\pi^{\mathcal{E}}_{\perp,a}\}$ and we reuse $\bar{\boldsymbol{\pi}}^{\mathcal{E}}(\cdot,\cdot)$ with its meaning determined by its arguments. Because the $\{\boldsymbol{v}^{\mathcal{E}}(\kappa_r), \boldsymbol{x}_{\perp,a}(\kappa_r) : 1 \le a \le N^{\mathcal{E}}\}$ are locally related to the fixed $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ basis vectors by a rotation, the Jacobian of $\bar{\boldsymbol{\pi}}^{\mathcal{E}} \leftrightarrow (s, \Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp})$ is unity.

Because the $\{\boldsymbol{x}_{\perp,a}\}$ are orthogonal

$$
\int_{-\Delta}^{\Delta} d\Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp} = \prod_{a=1}^{N^{\mathcal{E}}} \int_{-\Delta}^{\Delta} d\Delta\pi^{\mathcal{E}}_{\perp,a}
$$

and noting that

$${}^{\mathcal{P}}p(s, \Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp}) = {}^{\mathcal{P}}p(\kappa_r, \bar{\boldsymbol{\pi}}^{\mathcal{E}}_0) \left| \frac{\partial(s, \Delta\bar{\boldsymbol{\pi}}^{\mathcal{E}}_{\perp})}{\partial(\kappa_r, \bar{\boldsymbol{\pi}}^{\mathcal{E}}_0)} \right|^{-1} \;, \tag{69}$$

the integral of the prior over an orthotope with sides of infinitesimal length 2Δ centered about the trajectory point $\bar{\boldsymbol{\pi}}^{\mathcal{E}}[\kappa_r(s), \mathbf{0}]$ is

---

[18] $\mathcal{P}^{-1}(\kappa_r) \cdot \bar{\pi}^{\mathcal{E}}$ propagates backwards along a projected Poisson trajectory starting from interior point $\bar{\pi}^{\mathcal{E}}$. This does not circle endlessly and so must intersect the constraint region boundary. The intersection must be unique because it is not possible for $\bar{\pi}^{\mathcal{E}}_0 = (\mathcal{P}^{\mathcal{E}})^{-1}(\kappa_r) \cdot \bar{\pi}^{\mathcal{E}}$, $\bar{\pi}^{\mathcal{E}\prime}_0 = (\mathcal{P}^{\mathcal{E}})^{-1}(\kappa_r\prime) \cdot \bar{\pi}^{\mathcal{E}}$ if $\bar{\pi}^{\mathcal{E}}_0 \neq \bar{\pi}^{\mathcal{E}\prime}_0$. Proof: Eqs. (63) imply the composition property $\mathcal{P}^{\mathcal{E}}(\kappa_r\prime) \cdot \mathcal{P}^{\mathcal{E}}(\kappa_r) = \mathcal{P}^{\mathcal{E}}(\kappa_r\prime + \kappa_r)$, which implies that $(\mathcal{P}^{E})^{-1}(\kappa_r\prime) \cdot \mathcal{P}^{\mathcal{E}}(\kappa_r) = \mathcal{P}(\kappa_r - \kappa_r\prime)$. Using this and, without loss of generality, assuming that $\kappa_r\prime < \kappa_r$, $\bar{\pi}^{\mathcal{E}}_0 = \bar{\pi}^{\mathcal{E}\prime}_0$ would imply that $\bar{\pi}^{\mathcal{E}\prime}_0 = \mathcal{P}^{\mathcal{E}}(\kappa_r - \kappa_r\prime) \cdot \bar{\pi}^{\mathcal{E}}_0 = \mathcal{P}^{\mathcal{E}}(\Delta\kappa_r) \cdot \bar{\pi}^{\mathcal{E}}_0$. However, the positivity of all the lower-diagonal elements of $\mathcal{P}^{\mathcal{E}}(\Delta\kappa_r)$ and the fact that $(\bar{\pi}^{\mathcal{E}}_0)_0 > 0$ implies that all components of $\mathcal{P}^{\mathcal{E}}(\Delta\kappa_r) \cdot \bar{\pi}^{\mathcal{E}}_0$ are positive, so $\mathcal{P}^{\mathcal{E}}(\Delta\kappa_r) \cdot \bar{\pi}^{\mathcal{E}}_0 \in \text{int}(\Pi^{E})$. In contrast, $\bar{\pi}^{\mathcal{E}\prime}_0 \in \partial\Pi^{\mathcal{E}}_0$, so the equality cannot be true.

[19] While the subspace spanned by the $\{x_{\perp,a}\}$ is well-defined, the individual basis vectors are only defined up to a rotation about $v^{\mathcal{E}}(\kappa_r)$. This freedom is unimportant; any choice will do.

Figure 3.11. continued

$$\int_{-\Delta}^{\Delta}\int_{s-\Delta}^{s+\Delta} {}^{\mathcal{P}}p(s', \Delta\bar{\pi}_{\perp}^{\mathcal{E}})\, ds'\, d\Delta\bar{\pi}_{\perp}^{\mathcal{E}} = \int_{-\Delta}^{\Delta}\int_{s-\Delta}^{s+\Delta} {}^{\mathcal{P}}p[\kappa_r(s'), \bar{\pi}_0^{\mathcal{E}}(s', \Delta\bar{\pi}_{\perp}^{\mathcal{E}})] \left|\frac{\partial(s', \Delta\bar{\pi}_{\perp}^{\mathcal{E}})}{\partial(\kappa_r, \bar{\pi}_0^{\mathcal{E}})}\right|^{-1} ds'\, d\Delta\bar{\pi}_{\perp}^{\mathcal{E}}$$

$$\propto \int_{-\Delta}^{\Delta}\int_{\kappa_r(s)-\Delta/[ds(\kappa_r)/d\kappa_r]}^{\kappa_r(s)+\Delta/[ds(\kappa_r)/d\kappa_r]} \delta(\bar{\pi}_0^{\mathcal{E}} - \epsilon_0)\, F[\kappa_r(s)]\, d\kappa_r\, d\bar{\pi}_0^{\mathcal{E}}$$

$$= 2\Delta\, F[\kappa_r(s)] \left[\frac{ds(\kappa_r)}{d\kappa_r}\right]^{-1}_{\kappa_r=\kappa_r(s)} ,$$

where Eqs. (65) and (68) imply the definition of $\bar{\pi}_0^{\mathcal{E}}(s, \Delta\bar{\pi}_{\perp}^{\mathcal{E}})$. Constancy of the integral as s varies implies that

$$F(\kappa_r) \propto \frac{ds(\kappa_r)}{d\kappa_r} . \tag{70}$$

b. The posterior ${}^{\mathcal{P}}p(s|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$*:* For the reasons discussed above, the experimental data directly limits s, not κ_r, so we use the posterior distribution over this variable when computing narrowest credible intervals.[20] We compute it by marginalizing the posterior distribution in the context of the Poisson model over $\bar{\pi}^{\mathcal{E}}$. In terms of $p(\bar{\pi}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$, the posterior distribution computed from Eq. (53), this is

$$\begin{aligned} {}^{\mathcal{P}}p(\bar{\pi}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}) &= \frac{p(\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}|\bar{\pi}^{\mathcal{E}})}{p(\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})}\, {}^{\mathcal{P}}p(\bar{\pi}^{\mathcal{E}}) \\ &= \frac{p(\bar{\pi}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})}{p(\bar{\pi}^{\mathcal{E}})}\, {}^{\mathcal{P}}p(\bar{\pi}^{\mathcal{E}}) , \end{aligned} \tag{71}$$

where $p(\bar{\pi}^{\mathcal{E}})$ is the (uniform) prior that was used in computing ${}^{\mathcal{P}}p(\bar{\pi}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$. Applying Eqs. (64), (66), (70), (71), and the uniformity of $p(\bar{\pi}^{\mathcal{E}})$ we get

[20] We will see that in this case κ_r is tightly limited by the data so, with no significant loss in accuracy, we could compute the narrowest credible interval in κ_r directly from ${}^{\mathcal{P}}p(\kappa_r|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$. However, the difference will be important in the two-population Poisson model discussed below so, for consistency, we use the precise method here as well.

Figure 3.11. continued

$$
\begin{aligned}
{}^{\mathcal{P}}p(s|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}) &= \int {}^{\mathcal{P}}p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})\,\delta[s-s(\bar{\boldsymbol{\pi}}^{\mathcal{E}})]\,d\bar{\boldsymbol{\pi}}^{\mathcal{E}} \qquad (72)\\
&= \int \frac{p(\bar{\boldsymbol{\pi}}^{\mathcal{E}}|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})}{p(\bar{\boldsymbol{\pi}}^{\mathcal{E}})}\,{}^{\mathcal{P}}p(\bar{\boldsymbol{\pi}}^{\mathcal{E}})\,\delta[s-s(\bar{\boldsymbol{\pi}}^{\mathcal{E}})]\,d\bar{\boldsymbol{\pi}}^{\mathcal{E}}\\
&= C\int\int_0^\infty p[\bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r,\bar{\boldsymbol{\pi}}_0^{\mathcal{E}})|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}]\,{}^{\mathcal{P}}p(\kappa_r,\bar{\boldsymbol{\pi}}_0^{\mathcal{E}})\,\delta[s-s(\kappa_r)]\,d\kappa_r\,d\bar{\boldsymbol{\pi}}_0^{\mathcal{E}}\\
&= C\int\int_0^\infty p[\bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r,\bar{\boldsymbol{\pi}}_0^{\mathcal{E}})|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}]\,\delta(\bar{\boldsymbol{\pi}}_0^{\mathcal{E}}-\boldsymbol{\epsilon}_0)\,\frac{ds(\kappa_r)}{d\kappa_r}\delta[\kappa_r-\kappa_r(s)]\left[\frac{ds(\kappa_r)}{d\kappa_r}\right]^{-1}d\kappa_r\,d\bar{\boldsymbol{\pi}}_0^{\mathcal{E}}\\
&= C\,p\{{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}[\kappa_r(s)]\,|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}\}\,, \qquad (73)
\end{aligned}
$$

where C is the normalizing constant and we have used $\delta[s-s(\kappa_r)]\,ds(\kappa_r)/d\kappa_r = \delta[\kappa_r-\kappa_r(s)]$ and the appropriate Jacobians in going from the second to the third line.

c. Credible interval in κ_r^{wt}: We compute the narrowest (highest posterior density) one-sigma (68.3%) credible interval in s, $s_-\le s\le s_+$, from

$$
\begin{aligned}
&\arg\min_{[s_+,s_-]}(s_+-s_-)\ \text{s.t.}\ 0.683 =\\
&\qquad C\int_{s_-}^{s_+} p\{{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}[\kappa_r(s)]|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}\}\,ds\\
&\qquad = C\int_{\kappa_r(s_-)}^{\kappa_r(s_+)} p[{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}]\,\frac{ds(\kappa_r)}{d\kappa_r}\,d\kappa_r\,.
\end{aligned}
$$

We then map the interval into κ_r to get the limits in the homogenous Poisson model ${}^{\mathcal{P}}\kappa_{r\,\pm}^{\mathsf{wt}} = \kappa_r(s_\pm)$, where $\kappa_r(s)$ is computed numerically from Eq. (67). This gives for the homogeneous Poisson model estimate[21]

$$
{}^{\mathcal{P}}\kappa_r^{\mathsf{wt}} = (1.57\pm 0.01)\,. \qquad (74)
$$

As per Eq. (60) this corresponds to a mean number of replications ${}^{\mathcal{P}}\langle n\rangle = 3.14$.

The corresponding mean and variances of the best-fit estimate of ${}^{\mathcal{P}}\bar{\boldsymbol{\pi}}^{\mathsf{wt}}$ are

$$
\begin{aligned}
E({}^{\mathcal{P}}\bar{\boldsymbol{\pi}}^{\mathcal{E}}) &= C\int_0^\infty p[{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}]\,{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)\,\frac{ds(\kappa_r)}{d\kappa_r}\,d\kappa_r\\
\sigma_n^2 &= C\int_0^\infty p[{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}]\,\left\{{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)-E({}^{\mathcal{P}}\bar{\boldsymbol{\pi}}^{\mathcal{E}})\right\}_n^2\,\frac{ds(\kappa_r)}{d\kappa_r}\,d\kappa_r\,.
\end{aligned}
$$

These values are compared with the experimental estimate of $\bar{\boldsymbol{\pi}}^{\mathsf{wt}}$ in Figure 2F. The KO values are compared in Figure 2G. [22] While the WT fit is very good, the KO fit between ${}^{\mathcal{P}}\bar{\boldsymbol{\pi}}^{\mathsf{ko}}$ and $\bar{\boldsymbol{\pi}}^{\mathsf{ko}}$ is far outside the error bars. We see that the KO cells have an excess of highly proliferative cells that can not be explained by homogeneous Poisson replication, even if time-varying cell replication and loss rates are allowed.

B. Two-population Poisson model

Figure 2G shows that an appropriate representation of KO cell proliferation must include additional cells that replicate more times during the time period. A simple extension of the model is to add a second

---

[21] The fractional error in estimating ${}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}$ is lower than the fractional errors in $\bar{\boldsymbol{\pi}}^{\mathsf{wt}}$ because $N^{\mathsf{wt}}+2=9$ is large. Therefore the errors in the π_n^{wt} are almost independent, and the standard deviation of the projection onto the trajectory ${}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r)$ is much reduced.

[22] The KO one-sigma interval in the context of the homogeneous Poisson model (used in Figure 2G) is ${}^{\mathcal{P}}\kappa_r^{\mathsf{ko}} = 1.62\pm 0.03$. However, this is meaningless because the poor fit indicates that the model is incorrect.

Figure 3.11. continued

population of cells having a higher time-averaged Poisson replication rate κ_r^{hi}. These initially comprise fraction f_0^{hi} of the total number of cells, but this fraction will grow with time because of their (on average) faster replication. The normal (i.e., WT-like) and highly replicating cells evolve independently following Eq. (56). Therefore, the total population at time t is described by

$$ {}^{2\mathcal{P}}\eta_n^{\mathsf{ko}}\left(t; k_r^{\mathsf{wt}}, k_l^{\mathsf{wt}}, k_r^{\mathsf{hi}}, k_l^{\mathsf{hi}}\right) = \eta_0 \left[(1-f_{t_0}^{\mathsf{hi}})\,\mathcal{P}(n; 2\kappa_r^{\mathsf{wt}})\, e^{\Delta\kappa^{\mathsf{wt}}} + f_{t_0}^{\mathsf{hi}}\,\mathcal{P}(n; 2\kappa_r^{\mathsf{hi}})\, e^{\Delta\kappa^{\mathsf{hi}}}\right] , \tag{75} $$

where the first (second) term corresponds to the WT (KO) cells, and $\Delta\kappa^{\mathsf{wt}}$ and $\Delta\kappa^{\mathsf{hi}}$ are defined analogously to Eq. (57). The total numbers of normal and highly replicating cells grow exponentially according to Eq. (59) so f_t^{hi}, the fraction of highly replicating cells at time t, and f_0^{hi} are related by

$$ \frac{f_{t_0}^{\mathsf{hi}}}{1-f_{t_0}^{\mathsf{hi}}} = \frac{f_t^{\mathsf{hi}}}{1-f_t^{\mathsf{hi}}}\, e^{-(\Delta\kappa^{\mathsf{hi}}-\Delta\kappa^{\mathsf{wt}})} . \tag{76} $$

Dividing Eq. (75) by the total number of cells $\eta_0[(1-f_{t_0}^{\mathsf{hi}})\exp(\Delta\kappa^{\mathsf{wt}}) + f_{t_0}^{\mathsf{hi}}\exp(\Delta\kappa^{\mathsf{hi}})]$, applying Eq. (76), and fixing κ_r^{wt} at ${}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}$ [the rate for the normal sub-population estimated in Eq. (74) from the experiments with WT cells], we get

$$ {}^{2\mathcal{P}}\boldsymbol{\pi}\left(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}\right) = (1-f_t^{\mathsf{hi}})\,{}^{\mathcal{P}}\boldsymbol{\pi}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}) + f_t^{\mathsf{hi}}\,{}^{\mathcal{P}}\boldsymbol{\pi}(\kappa_r^{\mathsf{hi}}) . \tag{77} $$

This simple form is a consequence of the independent Poisson statistics of the normal and highly replicating populations. As in the homogeneous case, the rates of cell loss do not appear in the solution. (They are only needed to relate $f_{t_0}^{\mathsf{hi}}$ to f_t^{hi}.) Thus, the model has only two adjustable parameters, κ_r^{hi} and f_t^{hi}.

1. *Statistical estimation of ${}^{2\mathcal{P}}\kappa_r^{\mathsf{hi}}$ and ${}^{2\mathcal{P}}f_t^{\mathsf{hi}}$*

In this case, the model-consistent proportion vectors comprise a two-dimensional hypersurface parameterized by κ_r^{hi} and f_t^{hi}. Extending the previous discussion, we define the projected solution

$$ {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t) \equiv U^{\mathcal{E}} \cdot {}^{2\mathcal{P}}\boldsymbol{\pi}(\kappa_r^{\mathsf{hi}}, f_t) $$

and introduce a prior over $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ that is uniform on the hypersurface in $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$-space. In this case we reparameterize the $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$-space in terms of f_t^{hi}, κ_r^{hi}, and $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}}$:

$$ \bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}}) = U^{\mathcal{E}} \cdot \left[(1-f_t^{\mathsf{hi}})\,{}^{\mathcal{P}}\boldsymbol{\pi}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}) + f_t^{\mathsf{hi}}\,\mathcal{P}(\kappa_r^{\mathsf{hi}}) \cdot \bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}}\right] , $$

where $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}}$ is an alternative boundary condition for the highly replicating cells that is analogous to $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}}$ in the homogeneous case. As before, this mapping is invertible and the solution-space is obtained by restricting the boundary condition variable $\bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}} \to \boldsymbol{\epsilon}_0$:

$$ \bar{\boldsymbol{\pi}}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}, \boldsymbol{\epsilon}_0) = {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) . $$

Similarly, we restrict to solutions satisfying Eq. (77) by using a prior of the form

$$ {}^{2\mathcal{P}}p(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}}) \propto \delta(\bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}} - \boldsymbol{\epsilon})\, F(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) , $$

where $F(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$ is determined by the requirement that the integral of ${}^{2\mathcal{P}}p(\kappa_r^{\mathsf{hi}}, \bar{\boldsymbol{\pi}}_0^{\mathcal{E}\,\mathsf{hi}}, f_t^{\mathsf{hi}})$ over an infinitesimal orthotope surrounding an infinitesimal section of the hypersurface be constant as the orthotope is moved while maintaining its orientation relative to the surface.

In this case we define the orthotope using

$$ \boldsymbol{v}_{\kappa}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) = \frac{\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})/\partial\kappa_r^{\mathsf{hi}}}{\left|\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})/\partial\kappa_r^{\mathsf{hi}}\right|} \tag{78} $$

Figure 3.11. continued

and a second vector $\boldsymbol{v}_{\perp}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$, which is the component of

$$\boldsymbol{v}_f^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) = \frac{\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})/\partial f_t^{\mathsf{hi}}}{\left|\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})/\partial f_t^{\mathsf{hi}}\right|} \tag{79}$$

that is orthonormal to $\boldsymbol{v}_k^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$, as basis vectors for the hypersurface tangent space. Recalling that the proportion vector contraint limits $\bar{\boldsymbol{\pi}}^{\mathcal{E}}$ to $N^{\mathcal{E}}+1$ degrees-of-freedom, there are only $N^{\mathcal{E}}-1$ other locally defined orthonormal basis vectors, $\boldsymbol{x}_{\perp,a}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$. We designate the coordinates associated with $\boldsymbol{v}_\kappa^{\mathcal{E}}$, $\boldsymbol{v}_f^{\mathcal{E}}$, and the $\boldsymbol{x}_{\perp,a}$ as s, Δu, and $\Delta\bar{\boldsymbol{\pi}}_{\perp}^{\mathcal{E}}$ and note that the Jacobian of $\bar{\boldsymbol{\pi}}^{\mathcal{E}} \leftrightarrow (s, \Delta u, \bar{\boldsymbol{\pi}}_{\perp}^{\mathcal{E}})$, the local mapping in the vicinity of the solution hyperplane, is unity since, as above, this mapping is a rotation. Generalizing the constancy analysis we get

$$F(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) \propto \left|\frac{\partial(s, \Delta u)}{\partial(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})}\right| \ ,$$

Figure 3.11. continued

which leads to the posterior distribution[23]

$$
\begin{aligned}
&{}^{2\mathcal{P}}p(s,\Delta u|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}) = \\
&\qquad C_2\, p\{{}^{2\mathcal{P}}\pi^{\mathcal{E}}[\kappa_r^{\mathsf{hi}}(s), f_t^{\mathsf{hi}}(s,\Delta u)]|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}\}\ .
\end{aligned}
$$

a. Credible intervals in κ_r^{hi} and f^{hi}: Again we compute the narrowest one-sigma credible interval in κ_r^{hi} by mapping the interval in s. As discussed below, this indirect procedure is important in this case. The interval in s is determined from

$$
\begin{aligned}
\underset{[s_+,\,s_-]}{\arg\min}(s_+ - s_-) \text{ s.t.} \quad 0.683 &= \int_{s_-}^{s_+}\int {}^{2\mathcal{P}}p(s,\Delta u|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}})\, d\Delta u\, ds \\
&= C_2 \int_{s_-}^{s_+}\int p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}[\kappa_r^{\mathsf{hi}}(s), f_t^{\mathsf{hi}}(s,\Delta u)]|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}]\, d\Delta u\, ds \\
&= C_2 \int_{\kappa_r(s_-)}^{\kappa_r(s_+)}\int_0^1 p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\mathcal{X}^{\mathcal{E}},\mathcal{X}^{\mathcal{U}}] \left|\frac{\partial(s,\Delta u)}{\partial(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})}\right| df_t^{\mathsf{hi}}\, d\kappa_r^{\mathsf{hi}}\ . \qquad (80)
\end{aligned}
$$

The hyperplane Jacobian is

$$
\begin{aligned}
\left|\frac{\partial(s,\Delta u)}{\partial(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})}\right| &= \left|\frac{\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}\left(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}\right)}{\partial\kappa_r^{\mathsf{hi}}} \wedge \frac{\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}\left(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}\right)}{\partial f_t^{\mathsf{hi}}}\right| = \left| f_t^{\mathsf{hi}}\, \frac{\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}})}{\partial\kappa_r^{\mathsf{hi}}} \wedge \left[{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})\right]\right| \\
&= f_t^{\mathsf{hi}}\, \frac{ds(\kappa_r^{\mathsf{hi}})}{d\kappa_r^{\mathsf{hi}}}\, |{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})|\, |\sin\theta(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\ , \qquad (81)
\end{aligned}
$$

where $\wedge$ is the wedge product, which has the property that $|\boldsymbol{x}\wedge\boldsymbol{y}|$ is the area of the parallelogram defined by $\boldsymbol{x}$ and $\boldsymbol{y}$, $ds(\kappa_r^{\mathsf{hi}})/d\kappa_r^{\mathsf{hi}}$ is defined by Eq. (67) with the substitution $\kappa_r \to \kappa_r^{\mathsf{hi}}$, and $\theta(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$ is the angle between $\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}})/\partial\kappa_r^{\mathsf{hi}}$ and $[{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})]$. We simplify the Jacobian by exploiting the fact that the experimental $\bar{\boldsymbol{\pi}}^{\mathsf{ko}}$ differs from $\bar{\boldsymbol{\pi}}^{\mathsf{wt}}$ almost exclusively in magnitude of $\bar{\pi}_6$ relative to the division-resolved components $\{\bar{\pi}_0,\ldots,\bar{\pi}_5\}$ (cf Figures 2F and 2G). This implies that the posterior distribution will only have significant support where ${}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) \approx \boldsymbol{\epsilon}_6$ (i.e., where κ_r^{hi} is large). Performing an asymptotic expansion of Eq. (58) for $\kappa_r^{\mathsf{hi}} \to \infty$

$$
\begin{aligned}
{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) &\approx \boldsymbol{\epsilon}_6[1 - \mathcal{P}(5;2\kappa_r^{\mathsf{hi}})] + \boldsymbol{\epsilon}_5\mathcal{P}(5;2\kappa_r^{\mathsf{hi}}) \\
&= \boldsymbol{\epsilon}_6 + (\boldsymbol{\epsilon}_5 - \boldsymbol{\epsilon}_6)\frac{(2\kappa_r^{\mathsf{hi}})^5}{5!}e^{-2\kappa_r^{\mathsf{hi}}}\ . \qquad (82)
\end{aligned}
$$

In this approximation,

$$
\frac{\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}})}{\partial\kappa_r^{\mathsf{hi}}} \propto \boldsymbol{\epsilon}_6 - \boldsymbol{\epsilon}_5 \qquad (83)
$$

so to leading order

$$
\begin{aligned}
\cos\theta(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) &= \\
&\frac{\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}})/\partial\kappa_r^{\mathsf{hi}} \cdot [{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})]}{|\partial\, {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}})/\partial\kappa_r^{\mathsf{hi}}||{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})|} \\
&\approx \frac{(\boldsymbol{\epsilon}_6 - \boldsymbol{\epsilon}_5)\cdot[\boldsymbol{\epsilon}_6 - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})]}{|\boldsymbol{\epsilon}_6 - \boldsymbol{\epsilon}_5|\,|\boldsymbol{\epsilon}_6 - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})|}\ .
\end{aligned}
$$

Therefore, $|\sin\theta(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|$ is approximately constant and it, along with the term $|{}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}})|$ appearing in Eq. (81), can be absorbed into the normalizing constant. Then Eq. (80) simplifies to

---

[23] To be precise, we should also marginalize over ${}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}$, which appears in Eq. (77). However, because the credible interval in ${}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}$ is so small [Eq. (74)], this does not significantly affect the result so we simplify by holding ${}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}$ at its mean value.

Figure 3.11. continued

$$\underset{[s_+,\,s_-]}{\arg\min}(s_+ - s_-) \quad \text{s.t.} \quad 0.683 = C_2' \int_{\kappa_r(s_-)}^{\kappa_r(s_+)} \int_0^1 p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}]\, f_t^{\mathsf{hi}}\, \frac{ds(\kappa_r^{\mathsf{hi}})}{d\kappa_r^{\mathsf{hi}}}\, df_t^{\mathsf{hi}}\, d\kappa_r \,, \tag{84}$$

$$\text{where} \qquad C_2' = \left[\int_0^\infty \int_0^1 p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}]\, f_t^{\mathsf{hi}}\, \frac{ds(\kappa_r^{\mathsf{hi}})}{d\kappa_r^{\mathsf{hi}}}\, df_t^{\mathsf{hi}}\, d\kappa_r\right]^{-1} .$$

Applying the experimental data we find that the marginal posterior distribution ${}^{2\mathcal{P}}p(s|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}})$ has significant support as $s \to s(\infty)$, and this causes the solution to Eq. (84) to have $s_+ = s(\infty)$. Therefore, the credible interval mapped into κ_r^{hi} only specifies a lower bound for ${}^{2\mathcal{P}}\kappa_r^{\mathsf{hi}}$:

$$5.1 = \kappa_r^{\mathsf{hi}}(s_-) \leq {}^{2\mathcal{P}}\kappa_r^{\mathsf{hi}} < \infty \,. \tag{85}$$

(Note that if we had computed the narrowest credible interval in κ_r directly we would have obtained a finite upper bound for κ_r. This is incorrect: As discussed above, this only reflects the lack of sensitivity of the experiment, not a real experimental bound, and the bound is meaningless.) In agreement with the approximation of Eq. (82), ${}^{\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}[\kappa_r^{\mathsf{hi}}(s_-)] = 0.93\boldsymbol{\epsilon}_6 + 0.04\boldsymbol{\epsilon}_5 + \ldots$. That is, the approximation is accurate to within 3% at the lower bound and it is more accurate at larger values of κ_r^{hi}.

Eqs. (60) and (85) imply that the rapidly replicating subpopulation has divided on average more than ten times during PD21–47.[24]

To compute the credible interval in ${}^{2\mathcal{P}}f_t^{\mathsf{hi}}$ we reverse the roles of the direction vectors defined in Eqs. (78) and (79): We use $\boldsymbol{v}_f^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$ as one of the tangent-space basis vectors and use the orthonormal component of $\boldsymbol{v}_k^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})$ as the other. Designating the corresponding coordinates as s_f and Δu_k, we could use a calculation analogous to Eq. (84) to compute the narrowest credible interval in s_f and then map this into f_t^{hi}. However,

$$\begin{aligned} \frac{ds_f}{df_t^{\mathsf{hi}}} &= \left|\frac{\partial\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})}{\partial f_t^{\mathsf{hi}}}\right| \\ &= {}^{\mathcal{P}}\boldsymbol{\pi}(\kappa_r^{\mathsf{hi}}) - {}^{\mathcal{P}}\boldsymbol{\pi}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}) \\ &\approx \boldsymbol{\epsilon}_6 - {}^{\mathcal{P}}\boldsymbol{\pi}({}^{\mathcal{P}}\kappa_r^{\mathsf{wt}}) \,, \end{aligned} \tag{86}$$

where we use the leading term in the asymptotic expansion of Eq. (82) in the last line. This implies that f_t^{hi} is a linear function of s_f. Therefore, we get the same result by directly computing the narrowest credible interval in f_t^{hi} using

$$\underset{[f_+,\,f_-]}{\arg\min}(f_+ - f_-) \quad \text{s.t.} \quad 0.683 = C_2' \int_0^\infty \int_{f_-}^{f_+} p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}]\, f_t^{\mathsf{hi}}\, \frac{ds(\kappa_r^{\mathsf{hi}})}{d\kappa_r^{\mathsf{hi}}}\, df_t^{\mathsf{hi}}\, d\kappa_r$$

and get $f_t^{\mathsf{hi}} = 0.40 \pm 0.06$.

The corresponding mean and variances of the best-fit estimate of ${}^{2\mathcal{P}}\bar{\boldsymbol{\pi}}^{\mathsf{hi}}$ are

$$\begin{aligned} E({}^{2\mathcal{P}}\bar{\boldsymbol{\pi}}^{E}) &= C_2' \int_0^\infty \int_0^1 p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}]\, {}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})\, f_t^{\mathsf{hi}}\, \frac{ds(\kappa_r^{\mathsf{hi}})}{d\kappa_r^{\mathsf{hi}}}\, df_t^{\mathsf{hi}}\, d\kappa_r \\ \sigma_n^2 &= C_2' \int_0^\infty p[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}})|\mathcal{X}^{\mathcal{E}}, \mathcal{X}^{\mathcal{U}}] \left[{}^{2\mathcal{P}}\boldsymbol{\pi}^{\mathcal{E}}(\kappa_r^{\mathsf{hi}}, f_t^{\mathsf{hi}}) - E({}^{2\mathcal{P}}\bar{\boldsymbol{\pi}}^{\mathcal{E}})\right]_n^2 f_t^{\mathsf{hi}}\, \frac{ds(\kappa_r)}{d\kappa_r}\, d\kappa_r \end{aligned}$$

These values are compared with the experimental estimate of $\bar{\boldsymbol{\pi}}^{\mathsf{ko}}$ in Figure S2B. The fit is good and shows that the data is consistent with the KO cells being a mixture of a population that replicates with the WT rate and an additional subpopulation that replicates at least three times more. The highly replicating subpopulation constitutes $\sim 40\%$ of the total population at PD47.

---

[24] Heuristically, note that if the mean number of replications were lower an increase in $\bar{\pi}_5^{\mathsf{ko}}$, in addition to the increase in $\bar{\pi}_6^{\mathsf{ko}}$, would have been observed.

Figure 3.11. continued

III. MODELING THE STAGES OF THE HAIR CYCLE

The PD21–47 replication data reflects the composition of changes occurring during early anagen, anagen, catagen and telogen (Figure 2B). The PD24–27 and PD36–42 H2B-GFP dilution data (Figure 2H,J), which span small intervals, and the BrdU data (Figure 2K), which were collected across different hair cycle phases, can be used to quantitatively estimate the WT growth rate during the hair cycle and how this is affected by p21 KO.

A. Quantitation of changes in division rate during the hair cycle

b. WT replication rates: Previous experiments have shown that most WT bulge cell replication occurs during early anagen and that the replication rate is subsequently reduced to a low level by catagen (7, 8). We previously (7) found[25] that $\langle n\rangle^{\mathsf{wt}}_{22-25}$, the mean number of divisions from PD22–25, is 1.2 ± 0.1, corresponding by Eq. (60) to $\kappa_r^{\mathsf{wt}}{}_{22-25} = 0.61 \pm 0.04$. Converting this to an average replication rate $\langle k_r\rangle$ by

$$\langle k_r\rangle = \kappa_r/\Delta t \ , \tag{87}$$

where $\Delta t = t_1 - t_0 - \delta t$ is the dilution interval including a small adjustment δt to account for doxycycline uptake and H2B-GFP mRNA degradation,[26] gives the estimate

$$\langle k_r\rangle^{\mathsf{wt}}_{22-25} = 0.22 \pm 0.02/\text{day} \ .$$

The PD24–27 H2B-GFP data of this study (Figure 2H) provides an overlapping estimate in early anagen. The data are well-fit by the homogeneous model, and using the analysis of Sec. II A 1 and Eq. (87) gives

$$\langle k_r\rangle^{\mathsf{wt}}{}_{24-27} = 0.21 \pm 0.01/\text{day} \ .$$

This close agreement is remarkable since the mouse strains used, while closely related, were not identical.

We can use the PD36–42 data (Figure 2J) to estimate the replication rate during catagen:

$$\langle k_r\rangle^{\mathsf{wt}}_{36-42} = 0.03 \pm 0.05/\text{day} \ .$$

Comparing the rates we conclude that the replication rate decreases about seven-fold from the end of early anagen to the beginning of catagen

Comparing the PD42 and PD47 BrdU data (Figure 2K) suggests that there is an additional two-fold decrease in replication rate from catagen to telogen. However, the change cannot be accurately quantitated because the PD42 error bars are large and the comparison depends on the assumption that there is little change in the number of cells per bulge between PD42 and PD47.

c. p21 KO replication rates: The PD24–27 data (Figure 2H) shows that there is no detectable difference between the KO and WT replicates rates during early anagen, while the PD36–42 data (Figure 2J) shows that the overall rate of KO cell replication during catagen is

$$\langle k_r\rangle^{\mathsf{ko}}_{36-42} = 0.07 \pm 0.01/\text{day} \ ,$$

which is 2.3 times the WT rate.

---

[25] Although Ref. (7) cautions that the analysis used there is only guaranteed to provide a lower bound on $\langle n\rangle$, for PD22–25 it provides a good estimate of the true value because the histogram is well-resolved. The more sophisticated analysis of this paper provides good estimates of $\langle n\rangle$ under more general conditions.

[26] Based on prior experiments, we estimate $\delta t = 8$ hr. It is possible that errors in this estimate may introduce an additional 5% error in absolute rates, but this is not important for the analyses.

Figure 3.11. continued

B. Modeling the effect of p21 KO: Faster replication or prolonged replication?

The PD21–47 analysis suggests that the KO bulge cells at PD47 comprise two distinct populations—one behaving like WT cells and one, constituting $\sim 40\%$ of the total, that has replicated at three times more than WT over the period. In addition, the shift (relative to WT) of KO cells from the PD21–35 $n = 3$ to $n \geq 4$ bins (Figure 2I) suggests that there is a distinct subpopulation of highly replicated cells that constitute $\sim 9\%$ of the KO population by PD35. In contrast, there is no detectable difference between the WT and KO PD24–27 histograms (Figures 2H and 3A), suggesting that few highly replicated cells are present during early anagen. The simplest explanation of these observations is that the p21 KO induces by PD35 the appearance of a small subpopulation of highly replicating "HI" cells and that its fraction of the total population increases from 9% to 40% because of faster replication.

The PD21–47 analysis shows that ${}^{\mathcal{P}}\kappa_r^{\mathsf{hi}}{}_{21-47} - {}^{2\mathcal{P}}\kappa_r^{\mathsf{wt}}{}_{21-47} \geq 3.5$, and Eq. (76) implies that this would induce at least a 20-fold increase in the HI fraction if the HI and WT loss rates were the same. In principle, this could be the case if only a very small number (i.e, $\leq 2\%$) of HI cells were present initially. However, if the increased level of caspase-3^+ staining observed at PD35 (Figure 2K) reflects increased apoptosis primarily of the HI subpopulation, the increased HI cell loss would partially offset the replication difference-induced amplification. In this case, the HI cells would initially constitute a larger fraction (though $\leq 9\%$) of the total population that was amplified to a lesser extent.

Two simple but antipodal hypotheses could explain the emergence of the HI population: 1) The HI cells are present throughout the hair cycle and have a replication rate that is always a constant factor times the WT rate—$\langle k_r \rangle^{\mathsf{hi}}(t) = c \langle k_r \rangle^{\mathsf{wt}}(t)$, where $c \geq 3$. In this model, the HI cells are still subject to the mechanism that suppresses WT growth after early anagen; their increased growth over the entire cycle results from their continually higher growth rate. Alternatively, 2) the p21 KO has no effect during early anagen but a small fraction of KO cells escape the WT replication rate suppression at the end of early anagen. These continue to replicate at a high rate throughout the cycle. In this model, the increased growth over the entire cycle results from an extended growth period. In either case, the factors that cause only some cells to respond to the p21 KO could be either stochastic or deterministic (e.g., location within the bulge) and are unknown. The second possibility conforms with the observation that p21 expression increases after anagen (Figure 1B), but the first possibility or intermediate hypothesis that involve differences both in replication rate ratio and in replication rate suppression cannot be excluded.

IV. REFERENCES

1. Waghmare SK, Bansal R, Lee J, Zhang YV, McDermitt DJ, Tumbar T (2008) Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells. *EMBO J* 27:1309-1320.
2. Tumbar T, Guasch G, Greco V, Blanpain C, Lowry WE, Rendl M, Fuchs E (2004) Defining the epithelial stem cell niche in skin. *Science* 303:359-363.
3. Parks DR, Roederer M, Moore WA (2006) A new logicle display method avoids deceptive effects of logarithmic scaling for low signals and compensated data. *Cytometry A* 69A:541-551.
4. Bishop CM (2006) Pattern recognition and machine learning. (New York: Springer).
5. Connor RJ, Mosimann JE (1969) Concepts of independence for proportions with a generalization of the Dirichlet distribution. *J Amer Statist Assoc* 64:194-206.
6. Wong T (1998) Generalized Dirichlet distribution in Bayesian analysis. *App Math Computation* 97:165-181.
7. Zhang YB, White B, Shalloway D, Tumbar T (2010) Stem cell dynamics in mouse hair follicles: a story from cell division counting and single cell lineage tracing *Cell Cycle* 9: 1504-1510.
8. Tumbar T (2012) Ontogeny and homeostasis of adult epithelial skin stem cells. *Stem Cell Rev* 8:561-576.

REFERENCES

Attias, H. (1999). “Advances in Neural Information Processing Systems 12.” eds Solla, S.A., Leen, T.K., Muller, K-R (MIT Press, Cambridge, MA), pp 209–215.

Bishop, C.M. (2006). “Pattern Recognition and Machine Learning.” (Springer, New York).

Boedigheimer, M.J., Ferbas, J. (2008). “Mixture modeling approach to flow cytometry data.” Cytometry A **73**(5):421–429.

Cheng, T., N. Rodrigues, et al. (2000). "Hematopoietic stem cell quiescence maintained by p21cip1/waf1." Science 287(5459): 1804-1808.

Choudhury, A. R., Z. Ju, et al. (2007). "Cdkn1a deletion improves stem cell function and lifespan of mice with dysfunctional telomeres without accelerating cancer formation." Nat Genet 39(1): 99-105.

Cotsarelis, G. (2006). "Epithelial stem cells: a folliculocentric view." J Invest Dermatol 126(7): 1459-1468.

de Nooij, J. C., M. A. Letendre, et al. (1996). "A cyclin-dependent kinase inhibitor, Dacapo, is necessary for timely exit from the cell cycle during Drosophila embryogenesis." Cell 87(7): 1237-1247.

Devgan, V., C. Mammucari, et al. (2005). "p21WAF1/Cip1 is a negative transcriptional regulator of Wnt4 expression downstream of Notch1 activation." Genes Dev 19(12): 1485-1495.

Ferrandiz, N., J. M. Caraballo, et al. (2012). "p21 as a Transcriptional Co-Repressor of S-Phase and Mitotic Control Genes." PLoS One 7(5): e37759.

Foudi, A., K. Hochedlinger, et al. (2009). "Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells." Nat Biotechnol 27(1): 84-90.

Fuchs, E. (2009). "The tortoise and the hair: slow-cycling cells in the stem cell race." Cell 137(5): 811-819.

Hennings, H., A. B. Glick, et al. (1993). "FVB/N mice: an inbred strain sensitive to the chemical induction of squamous cell carcinomas in the skin." Carcinogenesis 14(11): 2353-2358.

Hoi, C. S., S. E. Lee, et al. (2010). "Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin." Mol Cell Biol 30(10): 2518-2536.

Hsu, Y. C., H. A. Pasolli, et al. (2011). "Dynamics between stem cells, niche, and progeny in the hair follicle." Cell 144(1): 92-105.

Kemp, C. J. (2005). "Multistep skin cancer in mice as a model to study the evolution of cancer cells." Semin Cancer Biol 15(6): 460-473.

Lapouge, G., K. K. Youssef, et al. (2011). "Identifying the cellular origin of squamous skin tumors." Proc Natl Acad Sci U S A 108(18): 7431-7436.

Marques-Torrejon, M. A., E. Porlan, et al. (2012). "Cyclin-Dependent Kinase Inhibitor p21 Controls Adult Neural Stem Cell Expansion by Regulating Sox2 Gene Expression." Cell Stem Cell.

Matsumoto, A., S. Takeishi, et al. (2011). "p57 is required for quiescence and maintenance of adult hematopoietic stem cells." Cell Stem Cell 9(3): 262-271.

Orford, K. W. and D. T. Scadden (2008). "Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation." Nat Rev Genet 9(2): 115-128.

Ortt, K. and S. Sinha (2010). "Chromatin immunoprecipitation for identifying transcription factor targets in keratinocytes." Methods Mol Biol 585: 159-170.

Osorio, K. M., S. E. Lee, et al. (2008). "Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation." Development 135(6): 1059-1068.

Osorio, K. M., K. C. Lilja, et al. (2011). "Runx1 modulates adult hair follicle stem cell emergence and maintenance from distinct embryonic skin compartments." J Cell Biol 193(1): 235-250.

Pippa, R., L. Espinosa, et al. (2011). "p27(Kip1) represses transcription by direct interaction with p130/E2F4 at the promoters of target genes." Oncogene.

Pyne, S., X. Hu, et al. (2009). "Automated high-dimensional flow cytometric data analysis." Proc Natl Acad Sci U S A 106(21): 8519-8524.

Scheitz, C. J., T. S. Lee, et al. (2012). "Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer." EMBO J 31(21): 4124-4139.

Speck, N. A. and D. G. Gilliland (2002). "Core-binding factors in haematopoiesis and leukaemia." Nat Rev Cancer 2(7): 502-513.

Suzuki, A., Y. Tsutomi, et al. (1998). "Resistance to Fas-mediated apoptosis: activation of caspase 3 is regulated by cell cycle regulator p21WAF1 and IAP gene family ILP." Oncogene 17(8): 931-939.

Tani, H., R. J. Morris, et al. (2000). "Enrichment for murine keratinocyte stem cells based on cell surface phenotype." Proc Natl Acad Sci U S A 97(20): 10960-10965.

Tesio, M. and A. Trumpp (2011). "Breaking the cell cycle of HSCs by p57 and friends." Cell Stem Cell 9(3): 187-192.

Topley, G. I., R. Okuyama, et al. (1999). "p21(WAF1/Cip1) functions as a suppressor of malignant skin tumor formation and a determinant of keratinocyte stem-cell potential." Proc Natl Acad Sci U S A 96(16): 9089-9094.

Trempus, C. S., R. J. Morris, et al. (2003). "Enrichment for living murine keratinocytes from the hair follicle bulge with the cell surface marker CD34." J Invest Dermatol 120(4): 501-511.

Tumbar, T. (2012). "Ontogeny and Homeostasis of Adult Epithelial Skin Stem Cells." Stem Cell Reviews and Reports 8(2): 561-576.

Tumbar, T., G. Guasch, et al. (2004). "Defining the epithelial stem cell niche in skin." Science 303(5656): 359-363.

van Os, R., L. M. Kamminga, et al. (2007). "A Limited role for p21Cip1/Waf1 in maintaining normal hematopoietic stem cell functioning." Stem Cells 25(4): 836-843.

Viale, A., F. De Franco, et al. (2009). "Cell-cycle restriction limits DNA damage and maintains self-renewal of leukaemia stem cells." Nature 457(7225): 51-56.

Waghmare, S. K., R. Bansal, et al. (2008). "Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells." EMBO J 27(9): 1309-1320.

Weinberg, W. C., E. Fernandez-Salas, et al. (1999). "Genetic deletion of p21WAF1 enhances papilloma formation but not malignant conversion in experimental mouse skin carcinogenesis." Cancer Res 59(9): 2050-2054.

White, A. C., K. Tran, et al. (2011). "Defining the origins of Ras/p53-mediated squamous cell carcinoma." Proc Natl Acad Sci U S A 108(18): 7425-7430.

Willis, S., C. L. Day, et al. (2003). "The Bcl-2-regulated apoptotic pathway." J Cell Sci 116(Pt 20): 4053-4056.

Yu, M., T. Mazor, et al. (2012). "Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors." Mol Cell 45(3): 330-343.

Zhang, Y. V., J. Cheong, et al. (2009). "Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells." Cell Stem Cell 5(3): 267-278.

Zhang, Y. V., B. S. White, et al. (2010). "Stem cell dynamics in mouse hair follicles: a story from cell division counting and single cell lineage tracing." Cell Cycle 9(8): 1504-1510.

Zhou, B. P., Y. Liao, et al. (2001). "Cytoplasmic localization of p21Cip1/WAF1 by Akt-induced phosphorylation in HER-2/neu-overexpressing cells." Nat Cell Biol 3(3): 245-252.

Zou, P., H. Yoshihara, et al. (2011). "p57(Kip2) and p27(Kip1) cooperate to maintain hematopoietic stem cell quiescence through interactions with Hsc70." Cell Stem Cell 9(3): 247-261.

CHAPTER 4

HAIR FOLLICLE STEM CELL QUIESCENCE IS ASSOCIATED WITH LOW LEVELS OF HISTONE H3K9ME3, H3K27ME3, AND H3K4ME3 MARKS AND HIGH GENOME PLASTICITY MEDIATED BY GROWTH FACTOR SIGNALING[3]

4.1 Introduction

Stem cells play a crucial role in maintaining adult tissues by giving rise to progenitors that further differentiate and form functional structures during homeostasis as well as injury. Loss of proper self-renewal and/or differentiation abilities of tissue stem cells may lead to failure of tissue maintenance, or may lead to cancer (Fuchs 2009; Daley 2012). Thus, understanding the molecular basis of how tissue stem cells balance between proliferation and quiescence, as well as how they control their fate to self-renew or differentiate, is of great importance in advancing stem cell biology and its applications to medicine.

Chromatin regulation has recently been shown to play crucial roles in stem cell biology (Orkin and Hochedlinger 2011). Embryonic stem cells possess hyperdynamic chromatin characteristic (Meshorer, Yellajoshula et al. 2006) along with overall low

[3] This chapter has not yet been published. Author contributions: K.C.Lilja assisted in experiments in Figure 4.2H-I and 4.6A; K.J.Colletier assisted in experiments in Figure 4.2H-I and 4.6B-D; L.Yao initiated and generated preliminary data for Figure 4.8A-E; L.Tu performed experiments and analyzed data for Figure 4.8F-K; C.S.Scheitz assisted in developing the quantitative ChIP-seq data analysis; J.Lee performed and analyzed all other experiments; J.Lee assembled all figures and wrote the first draft; J.Lee and T.Tumbar designed the research.

level of repressive histone modifications (Hawkins, Hon et al. 2010). Moreover, histone modifying enzymes act as a mediator of reprogramming (Onder, Kara et al. 2012) and repressive histone marks can act as a barrier during the reprogramming process (Chen, Liu et al. 2013). These suggest that global characteristics of histone modifications may influence embryonic stem cell genome plasticity, which become more restricted upon lineage commitment. However, whether similar processes take place in tissue stem cells as well as what regulates their global histone modification status remains unclear.

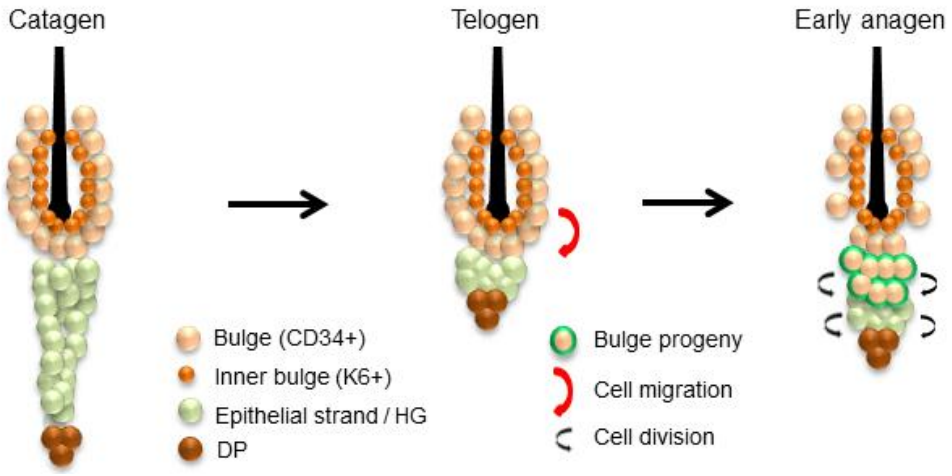
Hair follicle stem cells reside in a specialized niche called the bulge and undergo cyclic phases of proliferation and quiescence, which results in cycles of hair growth and regression respectively (Tumbar 2012). Prior to self-renewal some bulge cells migrate out from their niche in quiescence and become progenitors that further proliferate and differentiate into different lineages of hair follicle (Zhang, Cheong et al. 2009) (Figure 4.1). Later during the hair proliferative stage (anagen), bulge cells self-renew to make up for the loss. How these processes may be regulated via epigenetic mechanisms require further exploration.

Here we use mouse hair follicle as a model system to address the role of epigenetic regulation in balancing between proliferation and quiescence, as well as in fate determination. We show that multiple growth factors directly regulate the transcription of histone modifying enzymes, which bring about the globally low status of active (H3K4me3) and repressive (H3K9me3/H3K27me3) histone methylations, predisposing hair follicle stem cell genome into a plastic state. Coincidentally, the quiescent phase is the time of cell fate decisions in the hair follicle stem cells, where

one of two possible fates is chosen: (1) to migrate away from the niche, rapidly proliferate, terminally differentiate and eventually die by apoptosis, or (2) to remain as a stem cell in the niche and divide symmetrically to make more stem cells. We propose a model in which a robust epigenetic mechanism that controls histone methylation at specific marks couples quiescence and fate determination during normal tissue homeostasis.

Figure 4.1. Hair follicle stem cells during quiescence and fate determination.

Hair follicle stem cells enter quiescence in catagen. During quiescence, a few hair follicle stem cells migrate out from the native niche (bulge) and acquire progenitor cell fate. Later during the anagen, remaining cells in the bulge self-renew to replenish the stem cell pool (Zhang, Cheong et al. 2009). Image adopted and modified from (Lee and Tumbar 2012).


Catagen
Telogen
Early anagen
Bulge (CD34+)
Inner bulge (K6+)
Epithelial strand / HG
DP
Bulge progeny
Cell migration
Cell division

4.2 Materials and Methods

Mice, Growth factor-beads injections

All mice were treated according to Cornell University Institutional Animal Care and Use Committee protocols. We used CD1 x FVB wild-type strain for our analysis including histone methylation characterization and ChIP-seq. To induce reprogramming, we crossed K14-rtTA (from E. Fuchs, Rockefeller University, New York, NY) with *Col1a1-tetO-Pou5f1-Sox2-Klf4-Myc* (Jackson Laboratories B6;129S4-Col1a1$^{tm1(tetO\text{-}Pou5f1,\text{-}Klf4,\text{-}Sox2,\text{-}Myc)Hoch}$/J, #011001). For delivery of growth factor-coated beads, we used an approach previously described (Plikus, Mayer et al. 2008; Greco, Chen et al. 2009; Oshimori and Fuchs 2012). A small region of the mouse back skin was shaved. Recombinant mouse Noggin (R&D systems) was injected intradermally with FluoSpheres (Invitrogen) for 5 consecutive days on the same region. BrdU administration was done either intraperitoneally in PBS or via drinking water bottle (25μg/g).

Cell culture

For growth factor removal, mixed strain (129 x BL6) of wild-type newborn mouse skin was used to derive primary keratinocytes, which were expanded and used for experiments (between passages 4-6) (Osorio, Lilja et al. 2011). Cells were first grown to confluency. Then cells were washed once with PBS and added serum-free media containing either recombinant BMP4 (R&D systems) or its carrier only. Cells

incubated for 2, 4, 6, 8, 10, and 12 hours were harvested for RNA and histone protein extractions.

For fluorescence recovery after photobleaching (FRAP), primary keratinocytes were derived from *K5-rtTA* x *pTRE-H2BGFP*. Keratinocytes were treated with Doxycycline 2ug/mL for at least 48hrs to induce H2BGFP expression. Cells cultured in low Calcium media (0.05mM) were proliferating cells. Proliferating cells were treated with doxycycline for two days. Treatment with high Calcium (2mM) for 48hrs leads to keratinocyte differentiation. However, doxycycline induction is less efficient (fewer cells expressing H2B-GFP) under differentiation condition than cells under normal growth (potentially due to stratification). Therefore, cells were pre-treated with doxycycline for one day under low Calcium prior to two days incubation in high Calcium media. To keep cells in quiescent state, cells were starved by serial media dilution over three days (day 1: 50% of normal serum; day 2: 50% of normal serum; day 3: 25% of normal serum). Doxycycline was induced on the second day.

Immunofluorescence staining

Antibodies and dilutions used for staining are H3K4me3 (1:500; ActiveMotif), H3K9me3 (1:500; Millipore and Abcam), H3K27me3 (1:500-1:1000; Millipore), BrdU (1:300; Abcam), CD34 (1:150; BD Pharmingen). Protocol for staining have been previously described (Lee, Hoi et al. 2013). A fluorescence light microscope (Nikon) equipped with a charge-coupled device 12-bit digital camera (Retiga EXi; QImaging) was used to analyze all stained slides.

Western blot

For histone protein detection we FACS-isolated bulge and non-bulge cells. To extract histone proteins, cells were first incubated in Triton Extraction Buffer (PBS containing 0.5% Triton X 100 (v/v), 2mM phenylmethylsulfonyl fluoride), followed by 0.2N HCl incubation. H3K9me3 (1:500 – 1: 3000; Millipore) and H3 (1:3000 – 1: 10,000; Abcam) was used for immuoblotting.

Quantitative real-time PCR

A previously described method was used for RNA extraction and quantitative real-time PCR (Lee, Hoi et al. 2013). Primers used for qPCR are listed in Table 4.1.

Chromatin immunoprecipitation (ChIP) and ChIP-seq library preparation

For ChIP using keratinocytes, a previously described method was used (Lee, Hoi et al. 2013). 3-5 µg of H3 (Abcam), H3K4me3 (ActiveMotif), H3K9me3 (Millipore or Abcam), H3K27me3 (Millipore), and IgG (Cell Signaling) were used for immunoprecipitation. For ChIP using FACS-isolated cells, 15-20 mice of CD1 x FVB wild-type strain was used. ~3-4 million bulge and non-bulge cells nuclei were isolated and treated with micrococcal nuclease (MNase) (50U, Affymetrix) to generate mono-nucleosomal fragments. Isolated chromatin was processed further using described method (Lee, Hoi et al. 2013) for immunoprecipitation. Eluted ChIP DNA were either used for qPCR or further processed for sequencing library generation. ChIP-qPCR primers used are listed in Table 4.1. Library preparation was done using previously described method (Guertin and Lis 2010).

ChIP-seq data analysis

Single-end sequencing was performed using Illumina GAIIx (Catagen Non-bulge, 86nt) or HiSeq (Anagen bulge, 100nt). Initial analysis of ChIP-seq data were performed by the Computational Biology Service Unit (CBSU) of Cornell Bioinformatics Facility (Collaboration with Dr. Lishuang Shen). ChIP-seq reads were aligned using Bowtie (aligning tool) (Langmead, 2000) on mouse genome (mm9) and peaks were called using Model-Based Analysis of ChIPseq (MACS) (Zhang, Liu et al. 2008). However, MACS-called peaks are narrow and the peaks may not be directly comparable across different samples quantitatively. To resolve this problem, in consultation with Dr. Qi Sun, we are currently devising a new quantitative analysis pipeline. First, peaks will be called for each histone mark in each sample using MACS. Second, we will combine all peak locations called from all samples analyzed and create an annotation file. Thirdly, we will use a more quantitative RNA-seq pipeline known as 'Cuffdiff' (Gao, Kim et al. 2010; Kvam, Liu et al. 2012), which measures the number of reads (quantitative measurement of differential expression) sitting on the specific genomic location defined above and normalized per number of reads sequenced. For instance, if on certain genomic location, one sample has a peak but another sample does not, the comparison may not be quantitative. Thus, by generating a concatenated annotation file for all peaks called in all samples analyzed, we can feed the annotation file into Cuffdiff and quantitatively measure the number of reads on all genomic locations.

Fluorescence Recovery After Photobleaching (FRAP)

Tails of K5-tTA x pTRE-H2BGFP (Tumbar, Guasch et al. 2004; Lee, Hoi et al. 2013) mice were snipped, kept in PBS and brought to confocal microscopy immediately. Hair follicles at the edges of the tissue sample were imaged. A Zeiss confocal LSM 710, 30mW Argon/Neon laser at 80% power, 10 iterations was used to bleach 1/3-1/4 of the nucleus, including both heterochromatin and euchromatin. Images were collected at 1% power, every 10 seconds for 40 cycles, and processed with MetaMorph software. Data analysis and exponential curve fitting were performed by easyFRAP software (Rapsomaniki, Kotsantis et al. 2012). All statistical analyses were performed by two-sided student t-test. For FRAP recovery curves, a logarithmic trendline was fit for each curve of individual cells, and the slopes of the curves, representing how fast the recovery was, were subjected to the t-test.

Table 4.1. List of primers used for qPCR and ChIP-qPCR

Application	Gene/Region	Forward	Reverse
qPCR	Suv39h1	GCTCTGCCTCCTCTGAGGTAA	TCTCTGCATCTTCCGCACTA
qPCR	Suv39h2	CTGCCCAGGATAGCATTGTTC	CAAGTCTCGGCTCCACATTTAC
qPCR	Ezh2	AGTGACTTGGATTTTCCAGCAC	TCACCATGCACTTTTCCATCAT
qPCR	Jarid1a	TTTATCGGGCGCATCCGGCC	TCATTGCCTCAAGTTCATTCAGGCG
qPCR	Jmjd2a	GCGACCCTTGGTCTTCTTATT	GGAGAGCCGAGTGTGAAGAA
qPCR	Utx	CGGGCGGACAAAAGAAGAAC	CATAGACTTGCATCAGATCCTCC
qPCR	Setdb1	TGCAGCGTTTGTTACACTCA	CCAGCTAACTATCCAGGCCA
qPCR	Ehmt1	TGGCCACCACAAAGTCCCAGACA	GCTTTCTCGCCCTTGGCGCC
qPCR	Cbx3	ACTGGACCGTCGTGTAGTGAA	GCCCCTTGGTTTGTCAGCA
qPCR	Cbx5	GACAGGCGCATGGTTAAGG	CCTGGGCTTATTGTTTTCACCC
qPCR	Ezh1	TCCAGTATGTGATGCATTTGG	TGCGTCTCAGGATGGAGG
qPCR	Bmi1	TGATTCTGGTTGTTCGATGC	TGGCTCGCATTCATTTTATG
qPCR	Jmjd3	TGAAGAACGTCAAGTCCATTGTG	TCCCGCTGTACCTGACAGT
qPCR	Cbx6	CTCGGCTTGACAGAGAAATG	AGGGAACGTGAGCTGTATGG
qPCR	Ash1l	AAGTGCATGGTGTGGCAGCACTG	CTGGGCATAGTGGGGCCTAGGG
qPCR	Jarid1b	GAAGAGTTCGCGGACCCCTTCG	GTGTTTGGGCCTCCAGCTCATTCAG
qPCR	Jarid1c	CACTGTGGTAAGCACGAGGA	CTCGTCACCCTCATGAATCC
qPCR	Keratin1	AACCCGGACCCAAAACTTAG	CCGTGACTGGTCACTCTTCA
qPCR	Keratin5	CCTGCAGAAGGCCAAGCA	TGGTGTTCATGAGCTCCTGGTA
qPCR	CD34	AAGGCTGGGTGAAGACCCTTA	TGAATGGCCGTTTCTGGAAGT
qPCR	Oct4/Klf4 (transgenic)	TCTGTTCCCGTCACTGCTC	AGAGAGTTCCTCACGCCAAC
qPCR	Sox2/Myc (transgenic)	ACATGGGTAGTGGGCAATGT	CGCAGATGAAATAGGGCTGT
qPCR	Oct4 (endogenous)	CAAGGCAAGGGAGGTAGACA	CAAAATGATGAGTGACAGACAGG
qPCR	Klf4	ATGGTCAAGTTCCCAGCAAG	TTTGTAAGTCCGGGCATGTT

	(endogenous)		
qPCR	Sox2 (endogenous)	AGTGGTACGTTAGGCGCTTC	AAACCCAGCAAGAACCCTTT
qPCR	cMyc (endogenous)	TTGAGGAAACGACGAGAACA	GCCAAGGTTGTGAGGTTAGG
qPCR	Nanog (endogenous)	TACCTCAGCCTCCAGCAGAT	GATGCGTTCACCAGATAGCC
qPCR	Rex1 (endogenous)	CGGAGAGAAGCCGTATCAGT	CTTTGCGTGGGTTAGGATGT
qPCR	Dppa5 (endogenous)	CAGTCGCTGGTGCTGAAATA	GCCCGAATCTTGTTGTTTTG
qPCR	Fbx15 (endogenous)	GGCCTTGAATGGAGAACTGA	GCACACTCCTCCTACGCTGT
ChIP-qPCR	CD34	TCTCTCTCCTCCTGCCTTGA	AGCACTCCTTGTGCTTCTCC
ChIP-qPCR	Ltbp1	GCAGACAGTATCCGAGCAAA	CGCCAGGAAAAGACAAAGG
ChIP-qPCR	Bmp4	GCTCCCTTCTCCATAACCTG	ACTGCTGCCCAAACTGATG
ChIP-qPCR	Keratin1	CTCAGTATATAAGGGCACGGCACT	GACTCATGATGCCTTAGAGAGAGGT
ChIP-qPCR	Keratin10	AGCAAAGCCTAGCACCTGTGA	GCTGCTGGAGCTGTATAGAACAGAC
ChIP-qPCR	Major Satellite	GACGACTTGAAAAATGACGAAATC	CATATTCCAGGTCCTTCAGTGTGC
ChIP-qPCR	Minor Satellite	CATGGAAAATGATAAAAACC	CATCTAATATGTTCTACAGTGTGG

4.3 Results

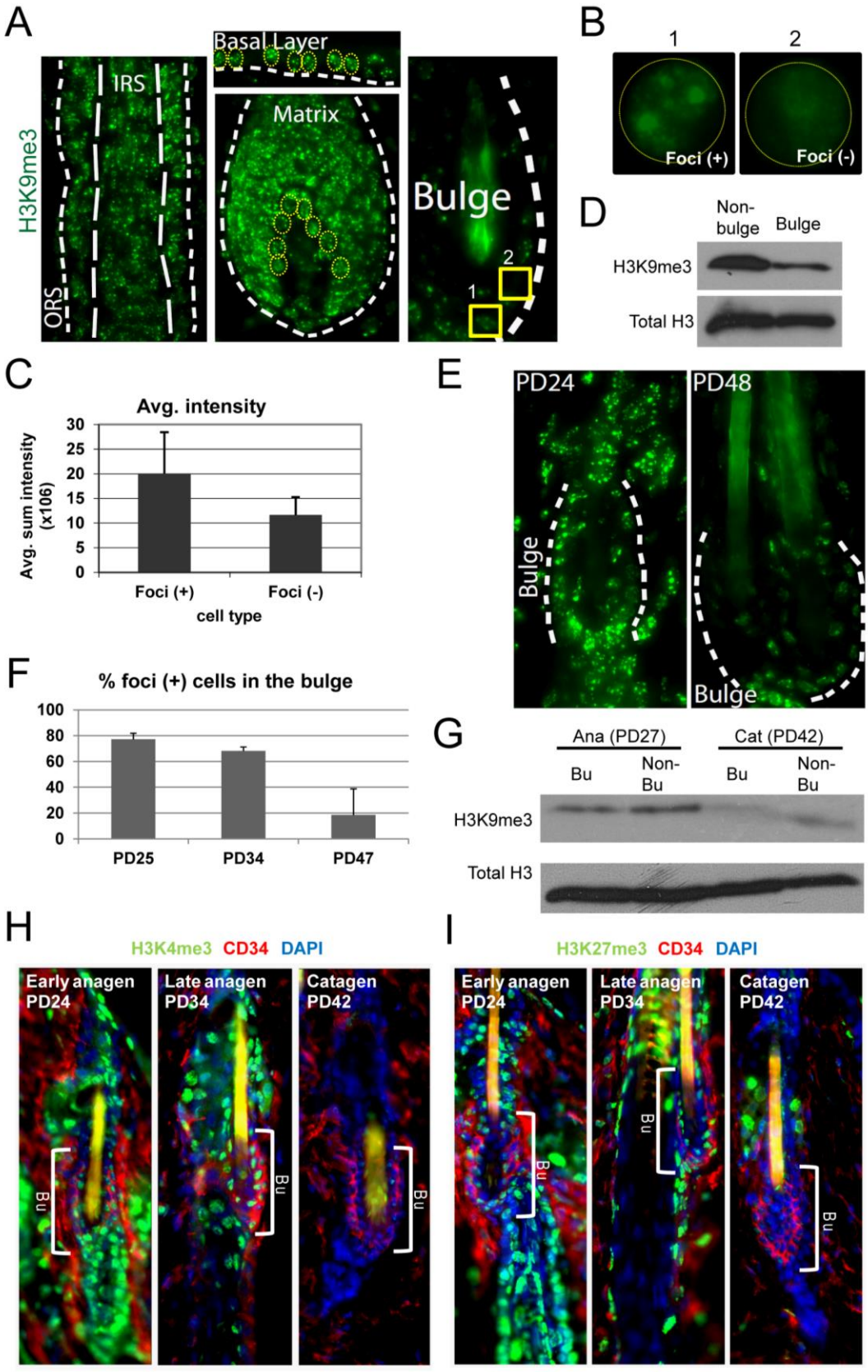
4.3.1 The status of H3 tri-methylation at K9, K27, and K4 in hair follicle stem cells is distinct and undergoes global erasure in quiescence

To begin to characterize the global chromatin status of hair follicle stem cells, we analyzed the level of H3K4me3 (active), H3K9me3 (repressive, heterochromatic), and H3K27me3 (repressive) histone modification marks by immunofluorescent staining of skin sections. First, H3K9me3 labeling revealed interesting cellular and global patterns. As expected, the H3K9me3 mark was universally present in different hair follicle/epidermal compartments including inner root sheath (IRS), outer root sheath (ORS), Matrix, and basal layer during anagen (Figure 4.2A). The H3K9me3 mark was dispersed throughout the nucleus with multiple distinct domains of heterochromatin. Strikingly, in the bulge, there were cells containing these heterochromatin domain-positive cells (foci+) and cells lacking these domains (foci-) (Figure 4.2B). Quantification of fluorescence intensities between these two cell types within the bulge, revealed a significant differences where foci- cells showed relatively lower signal than foci+ cells (Figure 4.2C). Moreover, western blotting of H3K9me3 using FACS (Fluorescence Activated Cell Sorting) isolated bulge (CD34+/α6-integrin+) and non-bulge (CD34-/α6-integrin+) cells indeed showed a significant differences in H3K9me3 levels (Figure 4.2D), suggesting that the bulge cells are globally more hypo-methylated than other hair follicle / epidermal compartments.

Next, we examined how the global status of histone methylation of hair follicle stem cells may change during the hair cycle. Strikingly, catagen/telogen bulge cells displayed significantly decreased global H3K9me3 level relative to anagen bulge

Figure 4.2. Hair follicle stem cells have distinct state of H3K9me3, H3K27me3, and H3K4me3, which undergoes global erasure upon quiescence

(A) H3K9me3 (green) staining of hair follicle in anagen. ORS: Outer Root Sheath; IRS: Inner Root Sheath. Notice bright heterochromatic positive (foci+) cells in yellow circles. Also note that the bulge contains two different localization patterns of heterochromatic 1) foci + and 2) foci -, which are enlarged in (B). (C) Quantification of fluorescence intensity of z-stack images of foci (+) and foci (-) cells (D) Western blot of H3K9me3 and total H3 of FACS-isolated bulge and non-bulge cells at anagen show clear differences in global H3K9me3 level (E) H3K9me3 localization pattern in anagen (PD24) and telogen (PD48). Note the clear qualitative differences in global level of H3K9me3 (F) Quantification of number of foci+ cells in the bulge across the hair cycle. (G) Western blot of H3K9me3 and total H3 of FACS-isolated bulge and non-bulge cells at anagen (PD27) and catagen (PD42) (H-I) Global localization pattern of H3K4me3 and H3K27me3 respectively during the hair cycle show similar behavior as H3K9me3.

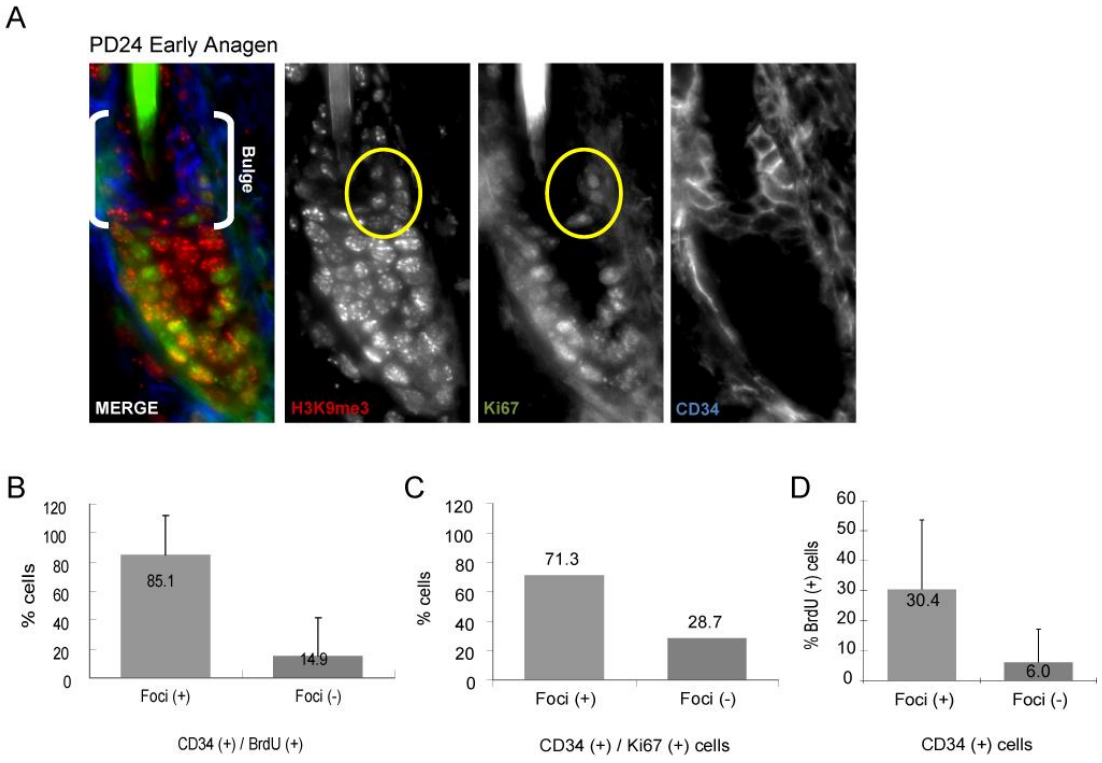

A
H3K9me3
IRS
ORS
Basal Layer
Matrix
Bulge
1
2
B
1
2
Foci (+)
Foci (-)
D
Non-bulge
Bulge
H3K9me3
Total H3
C
Avg. intensity
Avg. sum intensity (x106)
30
25
20
15
10
5
0
Foci (+)
Foci (-)
cell type
E
PD24
PD48
Bulge
Bulge
F
% foci (+) cells in the bulge
100
80
60
40
20
0
PD25
PD34
PD47
G
Ana (PD27)
Cat (PD42)
Bu
Non-Bu
Bu
Non-Bu
H3K9me3
Total H3
H
H3K4me3 CD34 DAPI
Early anagen PD24
Late anagen PD34
Catagen PD42
Bu
Bu
Bu
I
H3K27me3 CD34 DAPI
Early anagen PD24
Late anagen PD34
Catagen PD42
Bu
Bu
Bu

(Figure 4.2E). Indeed, quantification of number of foci+ cells as well as western blot using FACS-isolated bulge (CD34+/α6-integrin+) and non-bulge cells (CD34-/α6-integrin+) both revealed a significant reduction in global H3K9me3 levels as bulge cells enter quiescence (Figure 4.2E and F). Moreover, proliferating cells marked either by BrdU or Ki67 showed higher propensity of being foci+, suggesting that cells undergoing divisions gain H3K9me3, which then undergo significant erasure of this mark as these cells re-enter quiescence (Figure 4.3A-C). Notably, ~30% of bulge cells that are foci+ undergo proliferation, which is ~5-fold higher than foci- bulge cells (Figure 4.3D).

Interestingly, similar changes took place for H3K4me3 (active) and H3K27me3 (repressive) marks, suggesting that the global status of bulge cells histone H3 tri-methylations at different N-terminal lysine residues may be up-regulated during self-renewal but undergo global erasure upon re-entrance into quiescence (Figure 4.2H and I). A notable thing to mention was that global levels of H3K4me3 and H3K27me3 were also down-regulated to some extent in cells throughout the skin outside the bulge (Figure 4.2H and I). A previous study that employed a similar approach using mouse tail instead of dorsal skin did not document the quiescent versus proliferating follicle (Frye, Fisher et al. 2007). Collectively, these data suggest that hair follicle stem cells have overall low stats of both active (H3K4me3) and repressive (H3K9me3 /H3K27me3) histone marks that is distinct from other compartments of hair follicle and epidermis, which is increased during proliferation but undergoes global erasure upon quiescence. Presence of both active (H3K4me3) and repressive (H3K27me3) marks on a specific promoter in embryonic stem cells was considered a special status

Figure 4.3. Gain of H3K9me3 mark correlates with proliferation of bulge cells.

(A) Co-localization of H3K9me3, Ki67, and CD34 in early anagen shows that Ki67+ cells of the bulge gain H3K9me3 foci+ marks (B-C) Quantification of foci+ and foci- cells in CD34+/BrdU+ and CD34+/Ki67+ cells. Note that majority of dividing bulge cells are foci+ (D) Quantification of BrdU+ cells that are foci+ and foci- in CD34+ cells. Note the 5-fold increase in BrdU+ cells in foci+ relative to foci- cells.


A
PD24 Early Anagen
Bulge
MERGE
H3K9me3
Ki67
CD34
B
% cells
120
100
80
60
40
20
0
85.1
14.9
Foci (+)
Foci (-)
CD34 (+) / BrdU (+)
C
% cells
120
100
80
60
40
20
0
71.3
28.7
Foci (+)
Foci (-)
CD34 (+) / Ki67 (+) cells
D
% BrdU (+) cells
60
50
40
30
20
10
0
30.4
6.0
Foci (+)
Foci (-)
CD34 (+) cells

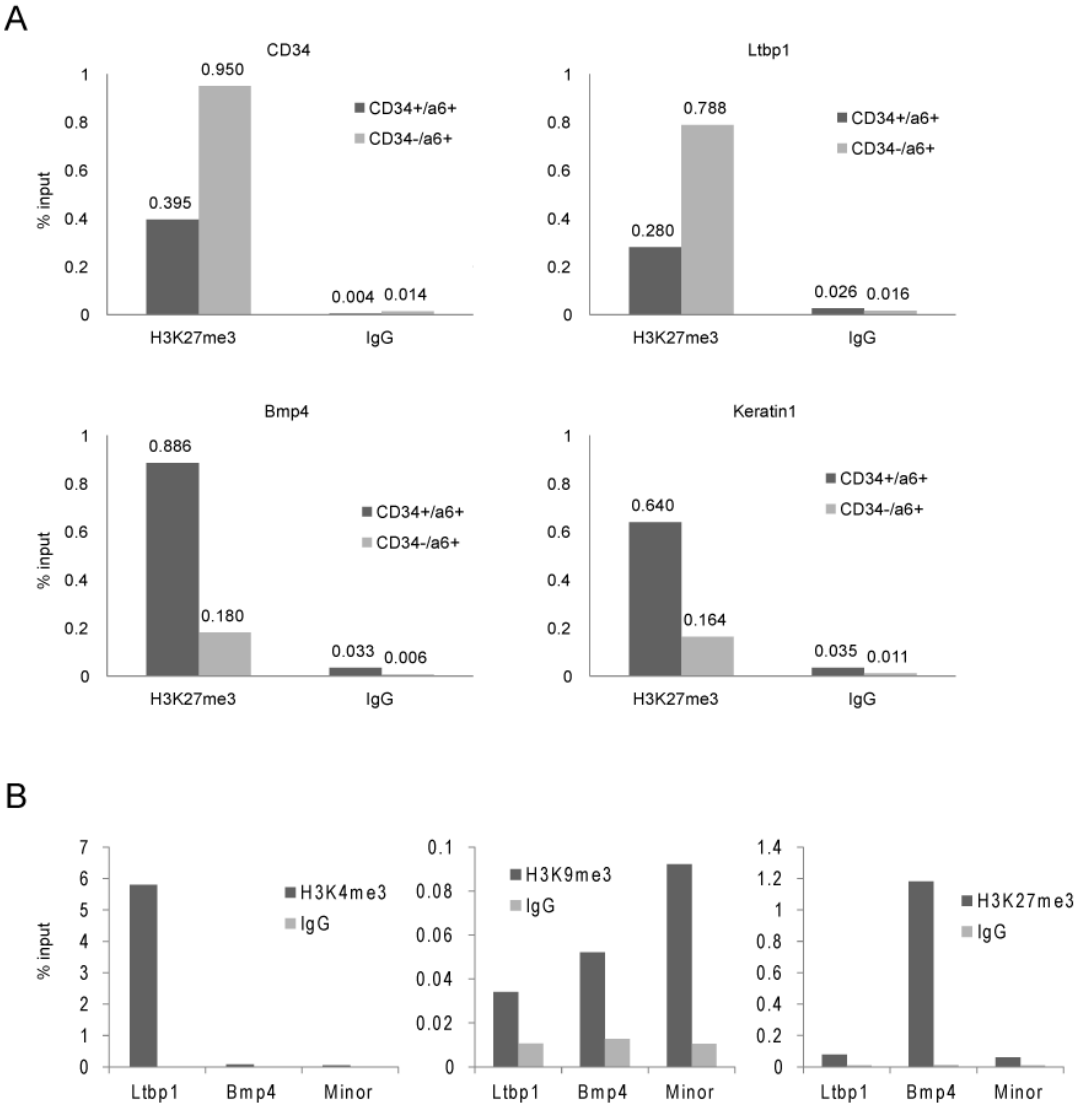
of bivalency that allows cells flexibility in either turning on and off specific lineage genes. We speculate that in tissue stem cells, erasure of both of these two marks as well as of H3K9me3 and potentially other marks may provide cells a state of low epigenetic identity which may confine high plasticity that coincides with the stage in which stem cells decide to either self-renew or differentiate.

4.3.2 Distinct genome-wide molecular changes upon hair follicle stem cells entrance to quiescence

Given the massive histone H3 demethylation that we observed in bulge cells at catagen, it is possible that these changes occurred not only in genic regions, but also in intergenic regions, non-genic regions, and repeat regions. Given that only a finite number of genes change expression upon quiescence such widespread changes in methylation would explain the obvious massive differences we observed at the cellular level by immunostaining and western blots. Next, we characterized in more depth the genomic localization of histone H3 tri-methylation on K9, K27, and K4 during hair follicle stem cells proliferation, quiescence, and differentiation using chromatin immunoprecipitation (ChIP) followed by deep sequencing (ChIP-seq) (Park 2009; Furey 2012). Although recently published work from another laboratory described a similar approach, the study never analyzed the status of H3K9me3, uses a different stage of quiescence (telogen), and did not use a quantitative approach to determine the level differences between different hair cycle stages (Lien, Guo et al. 2011). In fact,

Figure 4.4. Control experiments to validate the ChIP protocol using FACS-isolated or *in vitro* grown keratinocytes.

(A) ChIP-qPCR of FACS isolated bulge and non-bulge cells. CD34 and Ltbp1, which are expressed in the bulge show decrease in H3K27me3 in the bulge, but higher level in the non-bulge cells. On the other hand, Bmp4 and Keratin 1, which are expressed in non-bulge cells, display opposite pattern of H3K27me3. (B) Control ChIP-seq using *in vitro* keratinocytes show expected enrichments of different modification marks in specific regions.


A
CD34
% input
0.950
0.395
0.004
0.014
CD34+/a6+
CD34-/a6+
H3K27me3
IgG
Ltbp1
0.788
0.280
0.026
0.016
Bmp4
0.886
0.180
0.033
0.006
Keratin1
0.640
0.164
0.035
0.011
B
H3K4me3
H3K9me3
H3K27me3
Ltbp1
Bmp4
Minor

conventional analysis of histone modification distribution at promoters previously published did not parallel the changes in gene expression observed at telogen and anagen by microarray data (Lien, Guo et al. 2011). Finally, our immunostaining data suggest that by telogen the level of histone marks is elevated again, making catagen the stage with the lowest potential epigenetic identity as suggested by low methylation of critical transcriptional regulatory H3 marks (K9, K27, and K4) and this stage was never probed before by ChIP-seq.

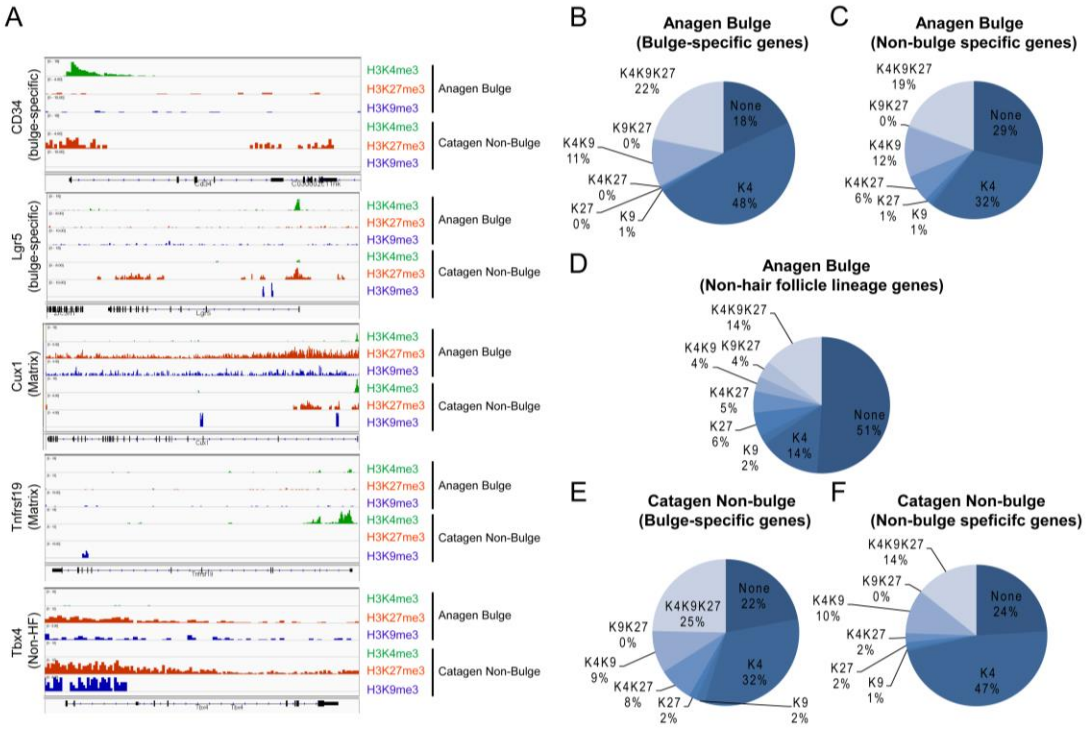
Thus, we FACS-isolated bulge (CD34+/α6-integrin+) and non-bulge (CD34-/α6-integrin+) cells from early anagen mouse skin when rapid self-renewal takes place with an increase in histone H3 tri-methylation marks (between PD22-25), and in catagen when the bulge cells are quiescent with globally low status of histone H3 tri-methylations (between PD39-43). Also, we are collaborating with CBSU bioinformaticians to develop a more quantitative approach using different pipelines (see Materials and Methods).

Due to the limitation in the number of sortable bulge cells per each mouse (~200,000 cells), we had to merge 15-20 mice per each stage to obtain enough cells for one experiment. Approximately 3-4 million sorted populations were split into five different ChIPs (H3, H3K4me3, H3K9me3, H3K27me3, and IgG). Experimental validation of H3K27me3 using FACS-isolated bulge and non-bulge cells by ChIP-qPCR clearly showed enrichments in expected regions, suggesting that our protocol is working (Figure 4.4A). In addition, validation of antibodies was done using cultured keratinocytes, where excess of cells can be obtained (Figure 4.4B).

Currently, we have sequenced DNA isolated by ChIp using antibodies against H3K4me3, H3K9me3, H3K27me3, and IgG from anagen bulge cells and catagen non-bulge cells. While our data are preliminary, our sequencing results do show clear expected differences in methylation patterns between the two populations (Figure 4.5). For instance, *CD34* and *Lgr5*, key bulge cell-specific markers, carry H3K4me3 in anagen bulge population, but carry H3K27me3 in catagen non-bulge, which fits with the mRNA expression patterns of these genes (Zhang, Cheong et al. 2009) (Figure 4.5A). On the other hand, genes expressed in the Matrix (differentiated) cells either show reduced H3K9me3/H3K27me3 (*Cux1*) or increased H3K4me3 (*Tnfrsf19*) in the catagen non-bulge. In addition, Tbx4, a non-hair follicle lineage gene, contains H3K9me3 and H3K27me3 on its promoter and gene body, which fits with its null expression in both populations (Figure 4.5A). Moreover, preliminary categorization of presence and/or absence of three marks within the list of genes specifically expressed in the bulge, in non-bulge, or outside the hair follicle lineages, revealed dramatic differences in composition (Figure 4.5B-F). For instance, in anagen, while ~48% of bulge-specific genes contained single H3K4me3, only ~32% of non-bulge genes and ~14% of non-hair follicle lineage genes contained single H3K4me3 (Figure 4.5B-D). In addition, in catagen non-bulge, while ~32% of bulge-specific genes carried single H3K4me3, this number increased up to ~47% in non-bulge-specific genes (Figure 4.5E and F). Although due to limiting starting material we could not perform serial ChIP for different histone marks, it is interesting to speculate on the existence of bivalent promoters (co-existence of H3K4me3 and H3K27me3 first observed in embryonic stem cells) (Bernstein, Mikkelsen et al. 2006; Mikkelsen, Ku et al. 2007),

Figure 4.5. Preliminary ChIP-seq data analyses reveal dramatic differences in different histone modifications in bulge and non-bulge cells.

(A) Snapshots of MACS-called peaks on different genomic locations from anagen bulge and catagen non-bulge using genome browser. (B-D) Distribution of histone marks (based on their presence or absence) in bulge-specific (B) (N=~300 genes), non-bulge-specific (C) (N=~500 genes), and not expressed in either populations (D) (N=~5000 genes) of anagen bulge population. (F-G) Distribution of histone marks in bulge-specific (E) (N=~300 genes) and non-bulge-specific (F) (N=~500 genes) of catagen non-bulge population. (K4: H3K4me3; K9: H3K9me3; K27: H3K27me3; K4K9: H3K4me3/H3K9me3; K4K27: H3K4me3/H3K27me3; K9K27: H3K9me3/H3K27me3; K4K9K27: H3K4me3/H3K9me3/H3K27me3).


A
CD34 (bulge-specific)
Lgr5 (bulge-specific)
Cux1 (Matrix)
Tnfrsf19 (Matrix)
Tbx4 (Non-HF)
H3K4me3
H3K27me3
H3K9me3
Anagen Bulge
Catagen Non-Bulge
B
Anagen Bulge
(Bulge-specific genes)
K4K9K27 22%
K9K27 0%
K4K9 11%
K4K27 0%
K27 0%
K9 1%
K4 48%
None 18%
C
Anagen Bulge
(Non-bulge specific genes)
K4K9K27 19%
K9K27 0%
K4K9 12%
K4K27 6%
K27 1%
K9 1%
K4 32%
None 29%
D
Anagen Bulge
(Non-hair follicle lineage genes)
K4K9K27 14%
K4K9 4%
K9K27 4%
K4K27 5%
K27 6%
K9 2%
K4 14%
None 51%
E
Catagen Non-bulge
(Bulge-specific genes)
K4K9K27 25%
None 22%
K9K27 0%
K4K9 9%
K4K27 8%
K27 2%
K9 2%
K4 32%
F
Catagen Non-bulge
(Non-bulge speficifc genes)
K4K9K27 14%
K9K27 0%
K4K9 10%
K4K27 2%
K27 2%
K9 1%
K4 47%
None 24%

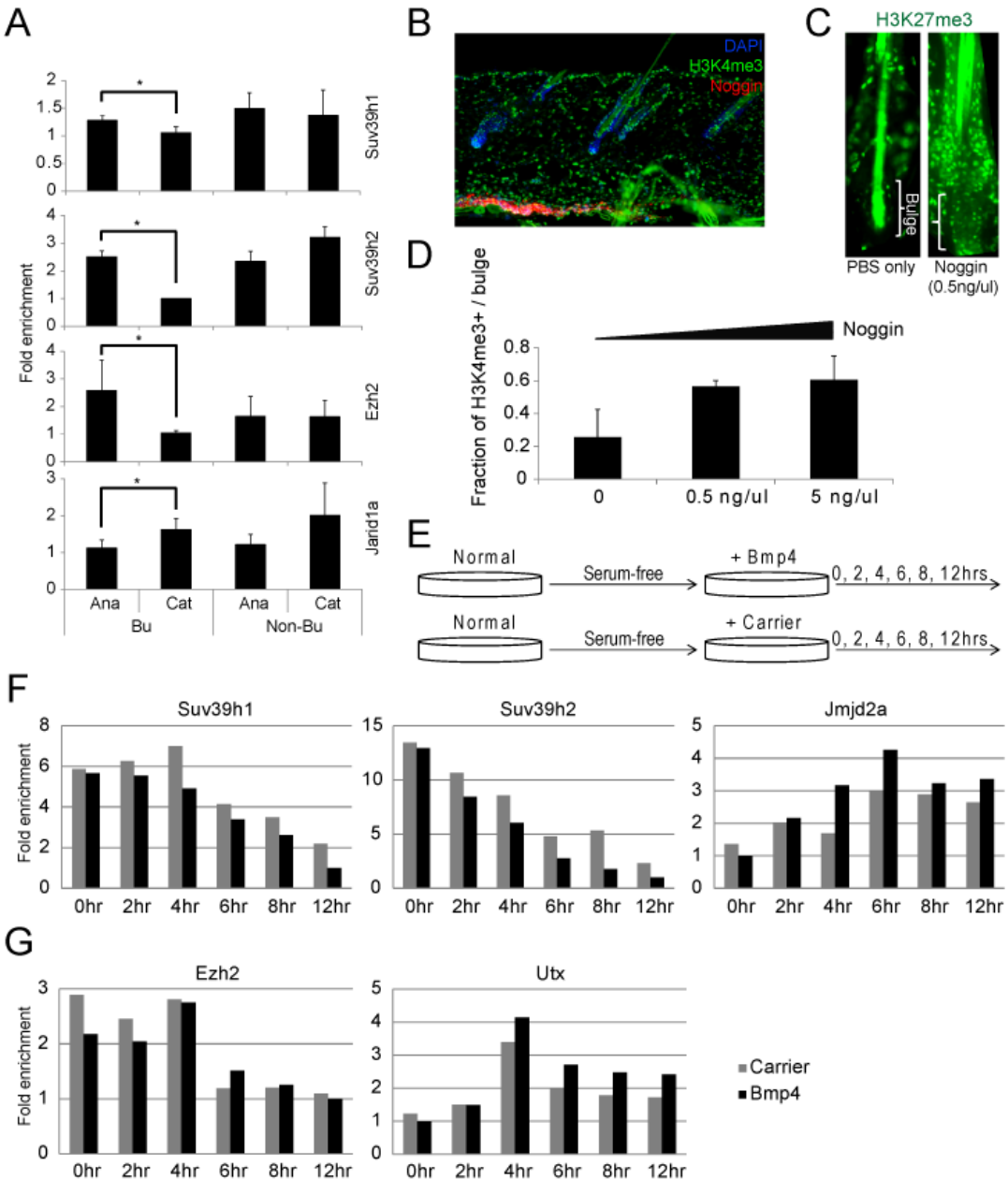
as suggested by their localization to the same promoters in the genome. Notably, this bivalency occurred in hair follicle cells often together with H3K9me3 (K4K9K27), and the proportions of what are dramatically different between different list of genes. It will be interesting in the future to elucidate if the presence and/or resolution of these bivalent domains may play regulatory roles during the hair follicle homeostasis (Figure 4.5B-F). Taken together, our preliminary analyses suggest that there seems to be correlation of active and/or repressive histone marks with gene expression patterns. These analyses are qualitative, as they are only based on the presence or absence of peaks in these marks relative to a general background level. Currently, more quantitative approach of calculating the differences in methylation levels across the genome is being developed using RNA-seq pipeline (see Materials and Methods for details). Using this approach, we expect to detect higher methylation levels across different genomic loci in proliferating bulge cells that are significantly reduced upon quiescence. In addition we expect to identify specific genomic loci of significant differences and further explore the biological meaning behind them.

4.3.3 Growth factor-mediated signaling regulates the global status of histone H3 tri-methylation at K9, K27, and K4 of hair follicle stem cells through direct transcriptional regulation of histone modifying enzymes

Next, we asked how different histone modifying enzymes that are responsible for placing or removing methylation marks are expressed during the proliferation and quiescence of hair follicle stem cells. Significant differences were observed from quantitative real-time PCR analysis of the mRNA expression level of these enzymes in

Figure 4.6. Growth factor-mediated signaling regulates the level of histone H3 tri-methylations of hair follicle stem cells through direct transcriptional regulation of histone modifying enzymes.

(A) Quantitative real-time PCR of specific histone modifying enzymes in FACS-isolated bulge and non-bulge cells from anagen and catagen. Note that the expressions of specific histone modifying enzymes differ and correlate with the global change in modification levels (B) Low magnification (10x) image of Noggin-soaked fluorescent beads-injected (red) catagen skin section, co-localized with H3K4me3 (green). Note the presence of red particles in dermal region in close proximity to the hair follicle. (C) H3K27me3 staining of PBS- or Noggin-injected hair follicle. Note the clear increase in global level of H3K27me3 upon Noggin injection. (D) Preliminary quantification of number of H3K4me3 positive cells per bulge in different Noggin treated skin sections. Note the clear increase in H3K4me3+ cells per bulge upon Noggin injection (E) Schematic of experimental strategy to address the effect of serum and Bmp4 in transcription of different histone modifying enzymes. (Carrier: solution used to dissolve Bmp4) (F-G) Quantitative real-time PCR of specific histone modifying enzymes in keratinocytes treated with serum-free (Carrier, grey bar) condition and serum-free + Bmp4 (black bar). Suv39h1, Suv39h2, and Jmjd2a are specific enzymes for H3K9me3, and Ezh2 and Utx are specific for H3K27me3. Note the decrease in methylases and increase in demethylases.


A
Fold enrichment
Suv39h1
Suv39h2
Ezh2
Jarid1a
Ana
Cat
Ana
Cat
Bu
Non-Bu
B
DAPI
H3K4me3
Noggin
C
H3K27me3
Bulge
PBS only
Noggin
(0.5ng/ul)
D
Fraction of H3K4me3+ / bulge
Noggin
0
0.5 ng/ul
5 ng/ul
E
Normal
Serum-free
+ Bmp4
0, 2, 4, 6, 8, 12hrs
Normal
Serum-free
+ Carrier
0, 2, 4, 6, 8, 12hrs
F
Suv39h1
Suv39h2
Jmjd2a
Fold enrichment
0hr
2hr
4hr
6hr
8hr
12hr
G
Ezh2
Utx
Fold enrichment
Carrier
Bmp4

FACS-isolated bulge (CD34+/α6-integrin+) and non-bulge (CD34-/α6-integrin+) cells (Figure 4.6A). For instance, Suv39h1/h2, two histone modifying enzymes specific for placing H3K9me3 (Rea, Eisenhaber et al. 2000; Lehnertz, Ueda et al. 2003), were specifically down-regulated in catagen bulge (CD34+/α6-integrin+) relative to anagen bulge (CD34+/α6-integrin+) (Figure 4.6A). Similar pattern was seen for Ezh2, a subunit of Polycomb Repressive Complex and a specific methylase for H3K27me3 (Margueron and Reinberg 2011). On the other hand, Jarid1a, a specific histone demethylase for H3K4me3, was up-regulated in catagen bulge relative to anagen bulge, correlating with the global erasure of H3K4me3 during quiescence (Figure 4.2). These data suggest that specific histone modifying enzymes are transcriptionally regulated during the hair cycle and their expression levels correlate with the global change in histone H3 methylation.

The transcriptional regulation of histone modifying enzymes as well as the global patterns of demethylation suggested that there are diffusible factors in the skin that might be at work. There is now a diverse study done in multiple signaling pathways mediated by intrinsic and/or extrinsic factors that play differential and important roles (Lee and Tumbar 2012). For instance, Bmp signaling is at high level during quiescence in the surroundings of the hair follicle including hair germ, inner bulge layer, dermal papillae and mesenchyme, and in adipocytes. On the other hand, Wnt and FGF (Fibroblast Growth Factor) signaling are up-regulated during proliferation. Inhibition of Bmp signaling as well as induction of Wnt signaling can activate hair follicle and induce proliferation. We reasoned that one or more of these

growth factors signaling might play important roles in inducing and/or removing global tri-methylation marks on histone H3.

To address this possibility, we used an inhibitor of BMP, Noggin, which we used to soak fluorescent beads and injected them into the back skin of catagen mice (Figure 4.6B) for five days (Plikus, Mayer et al. 2008; Greco, Chen et al. 2009; Oshimori and Fuchs 2012). We then sacrificed mice, and immunofluorescently labeled the skin sections for different histone H3 methylation marks. Strikingly, our preliminary results show that inhibition of Bmp signaling resulted in a significant increase in global methylation levels of H3K4me3 and H3K27me3 (Figure 4.6C and D), suggesting that Bmp signaling during quiescence inhibits (or erases) global methylation levels, thus maintaining the low epigenomic status of these histone modification marks.

Finally, we used *in vitro* keratinocytes to determine the effect of growth factors in expression of histone modifying enzymes. Keratinocytes *in vitro* have extensive proliferative ability, and the medium contains multiple growth factors (Barrandon and Green 1987). Strikingly, when serum was removed from the culture, expressions of histone modifying enzymes assessed by qPCR were significantly changed (Figure 4.6E). Preliminary results indicate that Suv39h1/h2 methylases were progressively down-regulated, while Jmjd2a, a K9-specific histone demethylase was progressively up-regulated over time (Figure 4.6F). In addition, Ezh2 was significantly down upon 6hr post-serum free while Utx, a K27-specific demethylase, was up-regulated by 4hrs post-serum free condition (Figure 4.6G). Moreover, when Bmp4 was added to the serum-free condition, the differences were enhanced further (Figure 4.6F and G),

suggesting that the transcriptional regulation of these histone modifying enzymes can be an additive effect of multiple signaling molecules including the Bmp.

Interestingly, some of these histone modifying enzymes carry conserved binding motifs for Smad4 (not shown), a co-transcription factors that works with other transcription factors upon Bmp signaling to regulate downstream gene expression. This observation suggested that the histone modifying enzymes we studied here might be directly regulated in their transcription by Bmp and other signaling pathways (Lee and Tumbar 2012). This hypothesis remains to be tested experimentally. Taken together, growth factors that promote or inhibit hair follicle stem cell quiescence regulate multiple histone modifying enzymes that are responsible for placing or removing specific genome-wide marks of histone H3 tri-methylations, potentially through direct transcriptional regulatory mechanisms.

4.3.4 Low level of histone H3 tri-methylation at K9, K27, and K4 marks in quiescence correlates with higher hair follicle stem cell genome plasticity for reprogramming

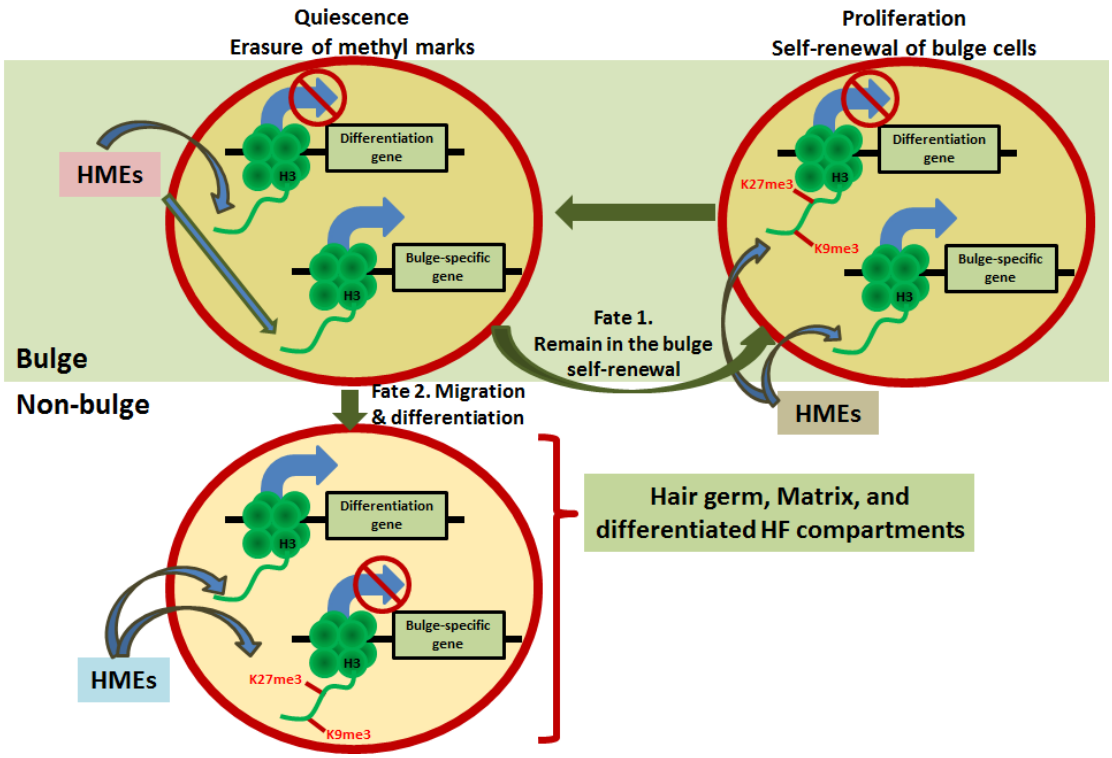
So far, we have shown that hair follicle stem cells at quiescence bear globally low tri-methylation of both active (H3K4me3) and repressive (H3K9me3/H3K27me3) histone H3 methylations, potentially regulated by direct transcriptional control mediated by multiple growth factors. Obvious questions would be the following: Why and how do these phenomena take place? What is the functional relevance? Interestingly, similar observations have been made in other systems. For instance,

embryonic stem cells epigenome contains distinct level of histone modification marks such as globally low H3K9me3 / H3K27me3, which are increased in more lineage restricted cell types (Hawkins, Hon et al. 2010). Also, chromatin of embryonic stem cells is more hyperdynamic than that of more lineage-restricted cell types observed by fluorescence recovery after photobleaching (Meshorer, Yellajoshula et al. 2006). In addition, quiescent B lymphocytes also contained globally low level of H3K9me3 and its associated proteins, which allowed these cells to be more efficiently reprogrammed upon nuclear transfer assays (Baxter, Sauer et al. 2004). Moreover, multiple histone modifying enzymes function in more efficient reprogramming of somatic cells into induced pluripotent cells (Onder, Kara et al. 2012) and H3K9me3 act as a barrier during reprogramming (Chen, Liu et al. 2013). Finally, hair follicle stem cells at quiescence choose their fate – either to migrate out and become differentiated progenitors or to stay in the bulge and later self-renew during anagen (Zhang, Cheong et al. 2009). Taken together, we hypothesized that the low level of histone H3 tri-methylation at K9, K27, and K4 is indicative of a potential globally low epigenomic status at quiescence that allows hair follicle stem cell genome to be more dynamic and plastic, allowing these cells to choose their fate (Figure 4.7).

To address this question, we took two different approaches. First, to ask whether the genome is generally more plastic at catagen, we induced reprogramming using the four core transcription factors (Takahashi and Yamanaka 2006) of both quiescent and proliferative bulge cells. The rationale here is that if the quiescent bulge cells have higher genome plasticity, the bulge cells would be more efficiently reprogrammed at catagen. Second, we used fluorescence recovery after

Figure 4.7. Model for epigenetic regulation of hair follicle stem cell quiescence and fate determination.

During quiescence, histone modifying enzymes are transcriptionally regulated by one or more growth factor signaling, leading to a global erasure of active and repressive histone methylation marks. The globally low epigenetic status of histone H3 tri-methylation in quiescent hair follicle stem cells puts the genome into a hyperdynamic state, which aids in fate choice. Prior to activation, some bulge cells can migrate out to hair germ where they are destined to differentiate (Fate 2), remaining bulge cells self-renew during anagen (Fate 1) to make up for the loss. HME: Histone Modifying Enzymes; Bulge-specific gene: genes specifically expressed in the bulge; Differentiation gene: genes expressed outside of the bulge.

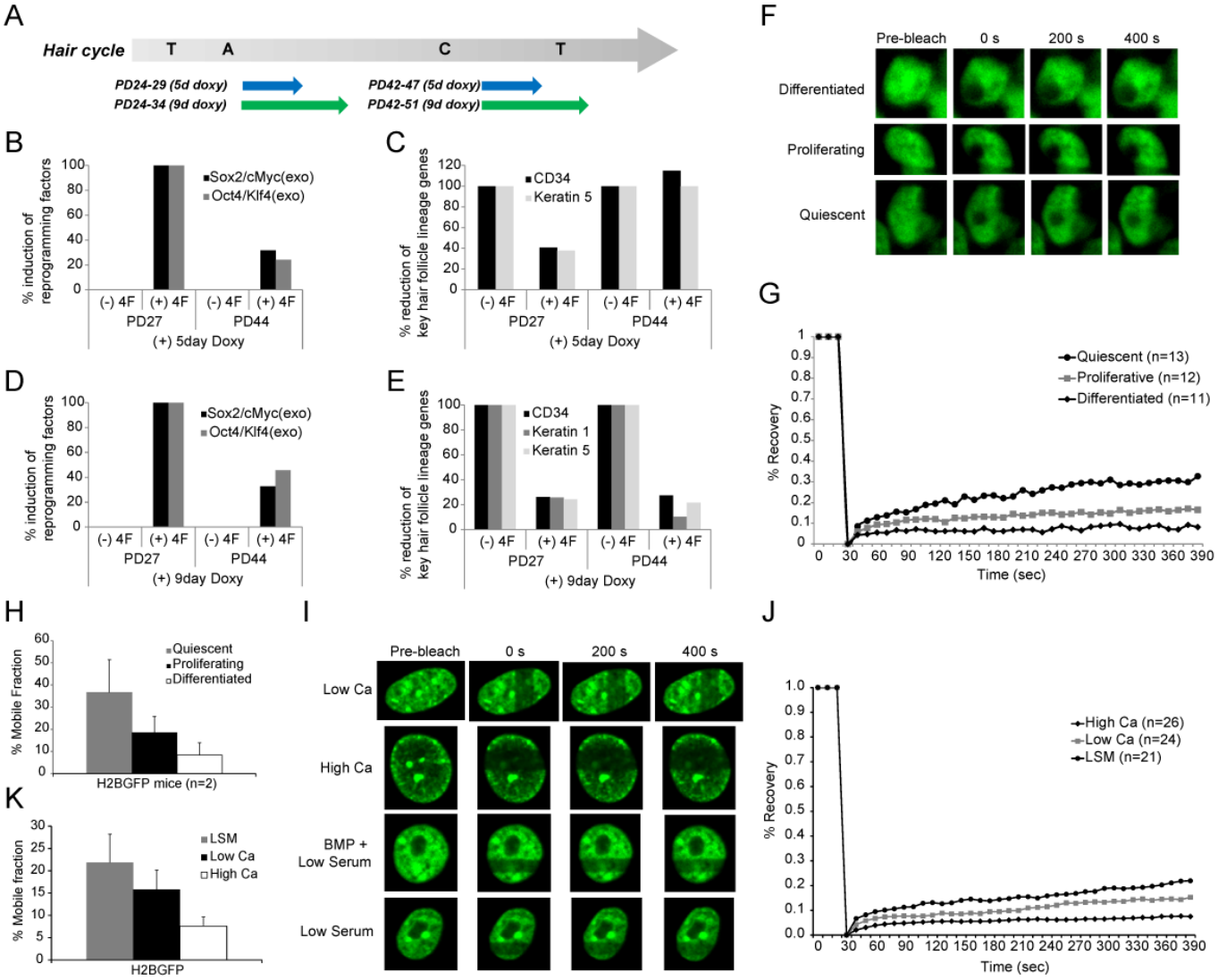

Quiescence
Erasure of methyl marks
Proliferation
Self-renewal of bulge cells
HMEs
Differentiation gene
Bulge-specific gene
H3
K27me3
K9me3
Bulge
Non-bulge
Fate 1.
Remain in the bulge
self-renewal
Fate 2. Migration
& differentiation
HMEs
Hair germ, Matrix, and
differentiated HF compartments
HMEs

photobleaching (FRAP) to assess the plasticity of the chromatin in proliferative and quiescent bulge cells.

First, to induce reprogramming, we crossed *K14-rtTA* doxycycline-inducible transgenic line with *Col1a-tetO-Oct4-Klf4-Sox2-cMyc* (*pTRE-OKSM*) transgenic line carrying inducible core reprogramming factors (Takahashi and Yamanaka 2006). We fed mice with doxycycline chow in anagen and catagen for 5 days and 9 days, sacrificed mice, and FACS-isolated bulge and non-bulge cells to analyze the expression of reprogramming factors (Figure 4.8A). Interestingly, while endogenous pluripotency factor Nanog was expressed by 9 days of induction in unsorted skin samples, none of the sorted populations expressed any of the endogenous reprogramming factors (data not shown). This may be due to the overall low reprogramming efficiency where only a small fraction of total cells are reprogrammed and they potentially lose CD34 expression at that time, and therefore cannot be sorted in the bulge fraction. Recent study suggests that only 5-10% of cells are permissive to reprogramming while 90-95% of cells are refractory (Polo, Anderssen et al. 2012). However, previous studies also showed that during the reprogramming process lineage-associated genes are down-regulated first and the identity of the originating cells is lost prior to expression of pluripotency markers (Stadtfeld, Maherali et al. 2008). Indeed we observed a significant reduction of hair follicle lineage-associated genes such as CD34 and Keratin5 upon 5 days of induction in anagen bulge (Figure 4.8B). However, same duration resulted in only ~20-30% of transgenes induction in catagen bulge, which was not sufficient to reduce CD34 and Keratin5 (Figure 4.8C). Strikingly, upon 9 day induction, while catagen bulge still expressed only ~30-40% of

Figure 4.8. Quiescent hair follicle stem cells are more hyperdynamic and are more readily reprogrammed.

(A) Experimental scheme of inducing reprogramming factors in anagen and catagen (B) Quantitative real-time PCR of 5-day induced transgenic reprogramming factors (each primer sets are designed to detect two transgenes) expressed in anagen (PD27) and catagen (PD44) bulge. Note that catagen bulge only express ~20-30% of transgenic reprogramming factors (normalized) relative to anagen bulge. (C) Quantitative real-time PCR of 5-day induced anagen and catagen bulge. Note the down-regulation of lineage-associated genes in anagen bulge but not in catagen bulge by 5 days (D) Quantitative real-time PCR of 9-day induced transgenic reprogramming factors show ~30-40% induction in catagen bulge (normalized) relative to anagen bulge. (E) Quantitative real-time PCR show similar extent of down-regulation of lineage-associated genes in anagen and catagen bulge. (F-H) Chromatin dynamics of H2B-GFP protein in hair follicle of K5tTA x pTRE-H2BGFP mice (n=2) (F) Representative image of fluorescence recovery after photobleaching (FRAP) (G) FRAP curves (H) Estimated mobile fractions in each condition after exponential curve fitting. Note that the quiescent bulge cells have the highest mobile fraction. (I-K) Chromatin dynamics of H2B-GFP protein in K5rtTA x pTRE-H2BGFP keratinocytes *in vitro* (I) Representative image of fluorescence recovery after photobleaching (FRAP). Low Ca: proliferating cells; High Ca: differentiated cells; LSM (low serum media): quiescent cells (J) FRAP curves (K) Estimated mobile fractions in each condition after exponential curve fitting. Note the highest mobile fraction in cells cultured under low serum media.


A
Hair cycle
T
A
C
T
PD24-29 (5d doxy)
PD24-34 (9d doxy)
PD42-47 (5d doxy)
PD42-51 (9d doxy)
B
% induction of reprogramming factors
Sox2/cMyc(exo)
Oct4/Klf4(exo)
(-) 4F
(+) 4F
PD27
PD44
(+) 5day Doxy
C
% reduction of key hair follicle lineage genes
CD34
Keratin 5
(+) 5day Doxy
D
(+) 9day Doxy
E
Keratin 1
(+) 9day Doxy
F
Pre-bleach
0 s
200 s
400 s
Differentiated
Proliferating
Quiescent
G
% Recovery
Quiescent (n=13)
Proliferative (n=12)
Differentiated (n=11)
Time (sec)
H
% Mobile Fraction
Quiescent
Proliferating
Differentiated
H2BGFP mice (n=2)
I
Low Ca
High Ca
BMP + Low Serum
Low Serum
J
High Ca (n=26)
Low Ca (n=24)
LSM (n=21)
K
% Mobile fraction
LSM
Low Ca
High Ca
H2BGFP

transgenes relative to anagen bulge, there was a similar extent of down-regulation of CD34, Keratin1 and Keratin5 (Figures 4.8D and E). These preliminary data potentially suggest that even with a lower amount of transgene induction, catagen bulge are more readily reprogrammed than the anagen bulge.

Secondly, using transgenic mice expressing H2B-GFP under repressible Keratin5 promoter (*K5tTA* x *pTRE-H2BGFP*) (Tumbar, Guasch et al. 2004), we employed fluorescence recovery after photobleaching (FRAP) to assess the dynamic nature of chromatin in quiescent and proliferative bulge cells as well as terminally differentiated cell types. Using whole tissue samples we photobleached H2B-GFP-expressing bulge cells in anagen (proliferating), in catagen (quiescent), and Matrix or Inner Root Sheath cells (differentiated) (Figure 4.8F-H). Strikingly, quiescent bulge cells showed the fastest recovery kinetics. H2B-GFP exchange rates, as well as estimated mobile fractions, were significantly lower in proliferating cells ($p<0.001$), followed by differentiated cells ($p<0.05$) (Figure 4.8F-H). In addition, FRAP analysis of *in vitro* keratinocytes expressing H2B-GFP (*K5rtTA* x *pTRE-H2BGFP*) cells cultured under different conditions (low Calcium: proliferating; high Calcium: differentiated; low serum media: quiescent) showed striking similarity to the *in vivo* data (Figure 4.8I-K). The recovery kinetics was fastest in quiescent cells, with the recovery rate of up to 22% of initial intensity (Figure 4.8I and J). Actively dividing cells recovered 15% ($p<0.05$) while differentiated cells recovered less than 10% ($p<0.0001$). One important control that we will add to these data is estimating the total amount of H2B-GFP in the different cell types we analyzed here. However, it is unlikely that especially in cell culture the rapid manipulation of conditions tested on

the same cell line may induce any detectable differences in overall H2B-GFP level. In addition, the rate of recovery is not dependent on the total amount of the protein but rather on the speed with which the protein move around. Our measurements here are not limited by normal diffusion, which has been measured on the order of seconds, while the recovery measured here is on the order of minutes. Additional studies will be required to consolidate these data. Taken together, these data may suggest that the quiescent bulge cells have higher genome plasticity as indicated by more productive early steps of reprogramming at this stage, and potentially a more dynamic chromatin with faster exchange rates of a core histone-histone H2B. We speculate that the high genome plasticity during quiescence potentially may contribute to bulge stem cells fate decision prior to activation.

4.4 Preliminary Conclusions

Here we use hair follicle as a model to decipher how the global status of important histone methylation marks of hair follicle stem cells is regulated. Although preliminary, our data suggest potential functional significance of this regulation during hair follicle homeostatic balance between proliferation, quiescence, and fate decision. Hair follicle stem cells appear to have a distinct epigenome characterized in part by globally low levels of both active (H3K4me3) and repressive (H3K9me3/H3K27me3) histone H3 marks relative to other hair follicle compartments (Figure 4.2). This status seems to be maintained by multiple growth factors signaling present around the hair follicle environment, which mediate direct transcription of responsible histone

modifying enzymes (Figure 4.6). Interestingly, the state of hair follicle stem cells quiescence correlates with a seemingly more dynamic chromatin and a more plastic state of their genome (Figure 4.8). Because this state coincides with the stage of hair follicle stem cells fate decision, it is tempting to speculate that it may enable the fate choice of quiescent hair follicle stem cells to decide either to migrate out from the niche and differentiate, or to remain in the niche and self-renew later to make up for the loss (Zhang, Cheong et al. 2009). Here we unveil a novel regulatory model through which growth factor mediated global histone H3 tri-methylation status at K9, K27, and K4 may couples the known state of tissue stem cell quiescence with that of low epigenetic identity or high plasticity, which is a hallmark of stem cell. Further work will be necessary to directly test several aspects of this model *in vivo*, and understand how widespread the erasure of histone marks, and potentially other components of the epigenome, is at tissue stem cell quiescence.

REFERENCES

Barrandon, Y. and H. Green (1987). "Three clonal types of keratinocyte with different capacities for multiplication." Proc Natl Acad Sci U S A 84(8): 2302-2306.

Baxter, J., S. Sauer, et al. (2004). "Histone hypomethylation is an indicator of epigenetic plasticity in quiescent lymphocytes." EMBO J 23(22): 4462-4472.

Bernstein, B. E., T. S. Mikkelsen, et al. (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell 125(2): 315-326.

Chen, J., H. Liu, et al. (2013). "H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs." Nat Genet 45(1): 34-42.

Daley, G. Q. (2012). "The promise and perils of stem cell therapeutics." Cell Stem Cell 10(6): 740-749.

Frye, M., A. G. Fisher, et al. (2007). "Epidermal stem cells are defined by global histone modifications that are altered by Myc-induced differentiation." PLoS One 2(8): e763.

Fuchs, E. (2009). "The tortoise and the hair: slow-cycling cells in the stem cell race." Cell 137(5): 811-819.

Furey, T. S. (2012). "ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions." Nat Rev Genet 13(12): 840-852.

Gao, D., J. Kim, et al. (2010). "A survey of statistical software for analysing RNA-seq data." Hum Genomics 5(1): 56-60.

Greco, V., T. Chen, et al. (2009). "A two-step mechanism for stem cell activation during hair regeneration." Cell Stem Cell 4(2): 155-169.

Guertin, M. J. and J. T. Lis (2010). "Chromatin landscape dictates HSF binding to target DNA elements." PLoS Genet 6(9).

Hawkins, R. D., G. C. Hon, et al. (2010). "Distinct epigenomic landscapes of pluripotent and lineage-committed human cells." Cell Stem Cell 6(5): 479-491.

Kvam, V. M., P. Liu, et al. (2012). "A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data." Am J Bot 99(2): 248-256.

Langmead, B. (2010). "Aligning short sequencing reads with Bowtie." Curr Protoc Bioinformatics Chapter 11: Unit 11 17.

Lee, J., C. S. Hoi, et al. (2013). "Runx1 and p21 synergistically limit the extent of hair follicle stem cell quiescence in vivo." Proc Natl Acad Sci U S A 110(12): 4634-4639.

Lee, J. and T. Tumbar (2012). "Hairy tale of signaling in hair follicle development and cycling." Semin Cell Dev Biol.

Lehnertz, B., Y. Ueda, et al. (2003). "Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin." Curr Biol 13(14): 1192-1200.

Lien, W. H., X. Guo, et al. (2011). "Genome-wide maps of histone modifications unwind in vivo chromatin states of the hair follicle lineage." Cell Stem Cell 9(3): 219-232.

Margueron, R. and D. Reinberg (2011). "The Polycomb complex PRC2 and its mark in life." Nature 469(7330): 343-349.

Meshorer, E., D. Yellajoshula, et al. (2006). "Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells." Dev Cell 10(1): 105-116.

Mikkelsen, T. S., M. Ku, et al. (2007). "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells." Nature 448(7153): 553-560.

Onder, T. T., N. Kara, et al. (2012). "Chromatin-modifying enzymes as modulators of reprogramming." Nature 483(7391): 598-602.

Orkin, S. H. and K. Hochedlinger (2011). "Chromatin connections to pluripotency and cellular reprogramming." Cell 145(6): 835-850.

Oshimori, N. and E. Fuchs (2012). "Paracrine TGF-beta signaling counterbalances BMP-mediated repression in hair follicle stem cell activation." Cell Stem Cell 10(1): 63-75.

Park, P. J. (2009). "ChIP-seq: advantages and challenges of a maturing technology." Nat Rev Genet 10(10): 669-680.

Plikus, M. V., J. A. Mayer, et al. (2008). "Cyclic dermal BMP signalling regulates stem cell activation during hair regeneration." Nature 451(7176): 340-344.

Polo, J. M., E. Anderssen, et al. (2012). "A molecular roadmap of reprogramming somatic cells into iPS cells." Cell 151(7): 1617-1632.

Rapsomaniki, M. A., P. Kotsantis, et al. (2012). "easyFRAP: an interactive, easy-to-use tool for qualitative and quantitative analysis of FRAP data." Bioinformatics 28(13): 1800-1801.

Rea, S., F. Eisenhaber, et al. (2000). "Regulation of chromatin structure by site-specific histone H3 methyltransferases." Nature 406(6796): 593-599.

Stadtfeld, M., N. Maherali, et al. (2008). "Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse." Cell Stem Cell 2(3): 230-240.

Takahashi, K. and S. Yamanaka (2006). "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." Cell 126(4): 663-676.

Tumbar, T. (2012). "Ontogeny and Homeostasis of Adult Epithelial Skin Stem Cells." Stem Cell Reviews and Reports 8(2): 561-576.

Zhang, Y., T. Liu, et al. (2008). "Model-based analysis of ChIP-Seq (MACS)." Genome Biol 9(9): R137.

Zhang, Y. V., J. Cheong, et al. (2009). "Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells." Cell Stem Cell 5(3): 267-278.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

5.1 Cell cycle control of hair follicle stem cell quiescence

Tissue stem cells are capable of proper maintenance of the tissue they reside in (Fuchs 2009). In order to do so, tissue stem cells adopt mechanisms to not only maintain their clones but also to produce enough progenitors that would differentiate and form the tissue structure. Therefore, even subtle changes in the cycling property of tissue stem cells could potentially be a threat to proper tissue homeostasis.

To begin to understand the cell cycle regulation during hair follicle cycling, we characterized how CDKis are expressed and regulated during the hair follicle homeostasis (Chapter 02). First, we found that multiple CDKis were differentially expressed (Figure 3.1). P15, p21, and p57 were up-regulated in the bulge during quiescence. In addition, p57 and p27 were always highly expressed in the bulge, presumably to maintain the overall low number of bulge cell divisions. The protein expression pattern of Runx1 inversely correlates with the CDKis expressions. Runx1 directly binds to the promoter of p21 and repress its transcription, and may potentially regulate p27 and p57 transcription as well.

To further dissect the role of CDKis, we coupled the previously developed cell division tracking system using pulse-chase of tet-repressible H2B-GFP protein

(Tumbar, Guasch et al. 2004; Waghmare, Bansal et al. 2008; Zhang, Cheong et al. 2009; Zhang, White et al. 2010), together with genetic ablation of p21. Interestingly, p21 knockout bulge cells failed to enter quiescence in catagen when wild-type bulge cells normally do (Figure 3.3 - 3.6). Instead p21 knockout bulge cells underwent 1-2 more rounds of divisions before entering quiescence. Extra bulge cells generated did not result in an increase in stem cell pool, but rather die more by apoptosis.

The significance of interaction between Runx1 and p21 was context-dependent. In keratinocytes, p21 deletion was sufficient to rescue the growth defect of Runx1 knockout (Hoi, Lee et al. 2010). However, in tumorigenic condition, p21 KO only partially rescued the defect in tumor formation of Runx1 knockout (Figure 3.7). More strikingly, *in vivo*, absence of both p21 and Runx1 resulted in further extension of the delayed activation. Intriguingly, double knockout of Runx1 and p21 *in vivo* and *in vitro* up-regulated p15 expression (Figure 3.9). Interestingly, we found that both Runx1 and p21 proteins directly bind to p15 promoter on their respective binding motifs, and co-work to repress p15 mRNA expression.

Together, these findings highlight several important perspectives of hair follicle biology as well as stem cell regulation in general. First, hair follicle stem cell quiescence is tightly regulated by multiple CDKis. During quiescence, p15, p21, p27, and p57 are up-regulated or maintained at high levels in the bulge. Considering the recent studies suggesting that one of the CDKis can compensate for the loss of another CDKi (Matsumoto, Takeishi et al. 2011; Tesio and Trumpp 2011; Zou, Yoshihara et al. 2011), it is plausible that multiple CDKis expressed in the bulge during quiescence

function redundantly. Our finding that p15 is up-regulated upon deletion of Runx1 and p21, also supports this possibility.

Secondly, p21 regulates the timely exit from the cell cycle during the normal hair follicle homeostasis. Using H2B-GFP division tracking system together with mathematical and computational modeling, we were able to map and predict the behavior of bulge cells division. On the contrary to a conventional idea of CDKis regulating the self-renewal rates, p21 regulates the extent of division of hair follicle bulge cells. Overall, the bulge cells undergo on average one extra round of division before arrest in catagen. Notably, knockout of p21 homolog (Dacapo) in Drosophila also results in one extra round of cell division before arrest (de Nooij, Letendre et al. 1996).

Third, Runx1, a key transcription factor of multiple stem cell systems (Speck and Gilliland 2002), directly regulate p21 transcription and potentially other CDKis (p27 and p57) as well. This suggests that out of many downstream targets that Runx1 acts on (Wilson, Foster et al. 2010), its action on CDKis expressions is an important axes of regulation that balance between stem cell proliferation and quiescence.

Fourth, we uncovered a novel role of p21 that is independent of its kinase inhibition function, in regulating p15 transcription. P21 together with Runx1 were bound to the p15 promoter to synergistically repress p15 mRNA expression. While recent studies suggested that two CDKis (p21 and p27) play cell-cycle independent, transcriptional repressive roles (Devgan, Mammucari et al. 2005; Pippa, Espinosa et al. 2011; Ferrandiz, Caraballo et al. 2012; Marques-Torrejon, Porlan et al. 2012), it was not previously known that one CDKi can directly regulate on the promoter of

another CDKis. This suggests that a robust mechanism might be at work to maintain stem cell quiescence while keeping the level of CDKis in stem cells in check. The latter might be crucial for rapid activation of stem cells in response to injury or specific tissue homeostatic signals. These ideas remain to be tested experimentally in the future.

Lastly, we unveil a robust multi-layered regulation of hair follicle stem cell quiescence. While Runx1 single knockout resulted in up-regulation of p21, p27, and p57, the levels of p27 and p57 appeared normal when p21 was knocked out in addition to Runx1. Given that p21 can bind to the p15 promoter and repress it, we can speculate that other CDKis may also work as CDKi transcriptional repressors. For instance, since double knockout t of Runx1 and p21 results in p15 up-regulation, p15 may now repress p27 and p57, explaining why the last two CDKis were not up-regulated in the Runx1/p21 double knockout. However, under Runx1 single knockout, p21 can maintain the repression of p15, which further allows p27 and p57 to be up-regulated. This idea, which we refer to as 'transcriptional circuitry of CDKis', needs further exploration, but offer a potential mechanisms to explain the robustness and multi-layered cell-cycle regulation in tissue stem cells to balance between proliferation and quiescence.

5.2 Regulation of hair follicle stem cell quiescence and genome plasticity via histone modifications

Epigenetics consists of broad range of regulatory mechanisms that are not governed by DNA sequence – DNA methylations, histone modifications, and other protein modifications (Kouzarides 2007). Of those, chromatin regulation is achieved through different histone modifying enzymes acting on different histone residues to place or remove chemical modifications that directly affect nearby transcription and/or openness of chromatin structure (Jenuwein and Allis 2001; Kouzarides 2007). Over the past decade, chromatin regulation of stem cells was shown to be a crucial mechanism that governs self-renewal and differentiation processes. However, if and how epigenetic mechanisms regulate hair follicle stem cells dynamic behavior of proliferation, quiescence, and differentiation, as well as how these regulation is achieved require further exploration.

To begin to understand the epigenetic regulatory mechanisms underlying the hair follicle homeostasis, we characterized global localization patterns of three histone H3 tri-methlation marks (Chapter 03). We discovered very interesting phenomena during this process. First, hair follicle stem cells possess globally lower levels of H3K4me3, H3K9me3, and H3K27me3 than other compartments of the hair follicle (Figure 4.2). Unexpectedly, while globally low, proliferating hair follicle bulge cells gained these modification marks to some extent, which was followed by a global reduction (erasure) upon quiescence (Figure 4.2). In addition, progress in genome-wide mapping using ChIP-seq is beginning to shed light on how hair follicle stem cells coordinate different marks of histone H3 tri-methylations in setting the state of proliferation, quiescence, and differentiation (Figure 4.5). Specific histone modifying enzymes for each mark such as Suv39h1, Suv39h2, Ezh2, and Jarid1a were

differentially expressed in the bulge at proliferation and quiescence. Interestingly, growth factor signaling including BMP signaling greatly influences the transcription of multiple histone modifying enzymes (Figure 4.6). We found that growth factors largely present during proliferative stage may bring increased levels of H3K4me3/H3K9me3/H3K27me3, whereas the absence of growth factors as well as the presence of inhibitory signaling such as BMP, leads to their reduction (erasure). Finally, we uncovered that this presumed low epigenomic status at quiescence potentially coincides with a more hyperdynamic chromatin state, which may potentially allow the more efficient reprogramming we have uncovered at this stage (Figure 4.8).

Our findings open up a new interesting perspective into understanding tissue stem cell biology and epigenetic regulations. First, growth factor-mediated signaling drives histone methylation changes. Although many studies revealed the existence and function of different histone modifying enzymes play specific roles in placement or removal of certain modification residues (Kouzarides 2007), relatively less was known about what drives the expression of these enzymes themselves. Multiple growth factors signaling such as Wnt and FGF are highly present around the stem cells during the growth phase of the hair follicle (Lee and Tumbar 2012). On the other hand, BMP signaling takes over during the quiescent phases of the hair cycle and is present in dermis, dermal papillae, inner bulge layer, and hair germ (Lee and Tumbar 2012). Interestingly, recent study revealed that BMP-signaling induces histone demethylases that removes repressive histone marks, leading to mesenchymal stem cell differentiation (Ye, Fan et al. 2012). This finding, together with our work, suggest that

downstream effectors of multiple growth factor signaling that are either intrinsic or surrounding the stem cells, may directly regulate the transcription of histone modifying enzymes. Intriguingly, histone modifying enzymes that were differentially expressed in the bulge at proliferation and quiescence such as Suv39h1, Ezh2 and Utx carry specific putative binding motifs for Smad4, a downstream effector of BMP signaling, which supports the notion that growth factor signaling may control the transcription of histone modifying enzymes.

Secondly, the status of histone H3 tri-methylation at K4, K9, and K27 changes in the transition between hair follicle stem cells proliferation and quiescence. During anagen, majority of bulge cells that were undergoing division gained H3K9me3 mark. More strikingly, Noggin induction in catagen to inhibit BMP signaling resulted in increased H3K4me3 and H3K27me3 marks without induction of proliferation. These findings suggest that histone methylation changes may occur prior to entering proliferation. This is consistent with previous studies showing that histone modifying enzymes action on cell cycle regulators affect the self-renewal ability of tissue stem cells (Fasano, Dimos et al. 2007; Ezhkova, Pasolli et al. 2009; Velichutina, Shaknovich et al. 2010; Ezhkova, Lien et al. 2011). In Chapter 2, we also showed that the p21 promoter shows increased enrichment of H3K27me3 upon Runx1 binding likely through recruitment of Polycomb Repressive Complex (Yu, Mazor et al. 2012). Also, recent studies in the cancer field showed that some of these histone modifying enzymes, and especially Ezh2, function to control cell cycle regulators in preventing metastasis (Jacobs, Kieboom et al. 1999; Molofsky, Slutsky et al. 2006; Bracken, Kleine-Kohlbrecher et al. 2007; Richter, Plehm et al. 2009; Velichutina, Shaknovich et

al. 2010). These studies suggest a hypothesis in which during quiescence, the removal of histone repressive marks on promoters such as p15, p21, and p27 results in up-regulation of their mRNAs, followed by cell cycle arrest at catagen. The quantitative analyses on more ChIP-seq data (presented in Chapter 4) will reveal this aspect. Taken together, these findings together suggest that epigenomic status arranged by histone modifying enzymes may be important in balancing between the proliferation and quiescence of stem cells.

Finally, recent studies in embryonic stem cells (Meshorer, Yellajoshula et al. 2006, Hawkins, Hon et al. 2010, Luger, Dechassa et al. 2012, Onder, Kara et al. 2012; Chen, Liu et al. 2013) as well as in more lineage restricted cell types (Baxter, Sauer et al. 2004) suggest that epigenomic status (as narrowly defined here by histone modifications) is known to influences stem cell genome plasticity. Our findings in the hair follicle stem cells seem to be consistent with these studies. Perhaps quiescent bulge cells with globally low epigenomic status in catagen may allow the chromatin to be more hyperdynamic under fluorescence recovery after photobleaching, be more plastic and readily reprogrammed. Most importantly, the genome plasticity during quiescence is likely to aid in fate decision of bulge cells prior to activation (Zhang, Cheong et al. 2009).

5.3 Perspective: unveiling the robustness of stem cell quiescence during hair follicle homeostasis

As discussed so far, tissue stem cells dynamic behavior during homeostasis is regulated by complex mechanisms and we are nowhere near the complete understanding of their biological implications. The two very different findings of this study – 1) the robust regulation of hair follicle stem cell quiescence through transcriptional control of cell cycle regulators, and 2) the correlation between growth factor-mediated low epigenetic status during hair follicle stem cell quiescence and high genome plasticity – give us some new insights.

Hair follicle stem cells in their niche remain relatively quiescent. Surrounding the niche, multiple growth factors that are activating and inhibitory signaling (Wnt, FGFs, BMPs and etc) from dermis, hair germ, dermal papilla, inner bulge layer, adipocytes, send signals to the niche to achieve the quiescence. These signals (such as BMP) can lead to activation/repression of histone modifying enzymes, which can lead to global change in epigenome. Often these histone modifying enzymes (such as Bmi1, Ezh2, Suv39h1) may co-work with intrinsic transcription factors (Runx1 and Nfatc1) to control the expressions of cell cycle regulators (p21 and p16) that are critical for proliferation. Hair follicle bulge cells have overall high level of multiple CDKis and even among these a complex transcriptional regulatory network may be at work to ensure that the proper level of CDKis are functioning to maintain quiescence.

Meanwhile, the globally low epigenomic status maintained by inhibitory growth factor signaling may bring the quiescent bulge cells chromatin to a more hyperdynamic and opened state. This may allow quiescent bulge cells to choose either to migrate out from the niche and acquire progenitor cell fate, or remain in the niche and participate in replenishing the stem cell pool. Importantly, our study couples the

quiescence, which is a common feature of multiple tissue stem cells such as hair follicle and blood, to high genome plasticity manifested by low epigenomic status. This connection, although has been implicated in progenitor cell types such as quiescent B lymphocytes (Baxter, Sauer et al. 2004), has not previously been described in tissue stem cells. Taken together, hair follicle stem cell quiescence appears to be robustly controlled by cell cycle regulators, and surrounding growth factors mediate genome-wide chromatin changes that may predispose the genome into an open chromatin state for assisting fate determination.

5.4 Future directions

While all contents of Chapter 3 have been published, data shown in Chapter 4 is preliminary. To strengthen our preliminary findings and make firm conclusions, further experiments must be pursued. Here I describe experimental details. First, I will perform ChIP-seq of missing populations – catagen bulge and anagen non-bulge – for H3K4me3, H3K9me3, and H3K27me3. IgG ChIP will be used as a control. Library preparation will be performed as (Guertin and Lis 2010). Sequencing will be done using single-end Illumina HiSeq available through Cornell Life Sciences Core Laboratories Center. Also, biological replicates will be necessary for all populations. All sequencing data will be analyzed as discussed in Materials and Methods section of Chapter 4. In addition, I will correlate the changes in histone H3 tri-methylation landscape with gene expression profiles previously obtained in our laboratory to assess

how these specific changes in histone H3 tri-methylations may influence gene expressions makeup of the self-renewing and quiescent bulge stem cells, as well as the makeup of the non-bulge cells. These analyses will bring quantitative insights into how genome-wide molecular epigenetic changes occur during the process of hair follicle stem cells proliferation, quiescence, and differentiation, and determine the role of these changes in cellular identities.

Second, I will address which growth factors are important for mediating the transcription of histone modifying enzymes *in vitro*. As shown in Figure 4.6, serum removal results in down-regulation of multiple histone modifying enzymes. Serum contains several growth factors that promote proliferation of keratinocytes. By using chemical inhibitors that inhibit specific growth factors or their binding receptors, I will be able to narrow down the list. Conversely, under serum-free condition, I can add recombinant growth factors that are important for *in vivo* activation of bulge cells such as FGF7, TGF-β2, Noggin, and Wnt molecules (Lee and Tumbar 2012) and analyze whether the expression of histone modifying enzymes will be reversed. After determination of specific growth factor(s) signaling, I will analyze the direct transcriptional regulation by chromatin immunoprecipitation followed by quantitative real-time PCR (ChIP-qPCR), to identify whether the downstream effector of the specific signaling binds to the promoter of different histone modifying enzymes. For instance, Suv39h1, Ezh2, Utx, and Jarid1a contain putative binding motifs for Smad4, a downstream effector of BMP signaling. In addition, to determine whether change in the expression of histone modifying enzymes indeed cause global change in the level of histone H3 tri-methylations, I will collect keratinocytes treated with normal serum,

serum-free, serum-free+Bmp4, and other chemical inhibitors or recombinant growth factors for histone protein extraction followed by western blotting of H3K4me3, H3K9me3, and H3K27me3. Taken together, these experiments will establish which growth factor(s) mediate the global changes is histone H3 methylation status by direct transcriptional regulation of specific histone modifying enzymes.

Third, I will further examine the functional relevance of quiescence, global low epigenomic status, and genome plasticity. Using inducible transgenic mice expressing core reprogramming factors (*K14-rtTA* x *Col1a1-tetO-Oct4-Klf4-Sox2-Myc*), I showed that even after 9 days of doxycycline treatment, neither the proliferating nor quiescent bulge cells expressed reprogramming / pluripotency markers (data not shown). This may be due to 1) insufficient duration of doxycycline treatment or 2) low efficiency of reprogrammed cells. Fibroblasts undergoing reprogramming pass through intermediate stages where they down-regulate lineage associated genes before expressing pluripotency markers suggesting that our bulge cells at 9 days doxycycline induction may be at an intermediate stage (Stadtfeld, Maherali et al. 2008; Polo, Anderssen et al. 2012). On the other hand, considering that reprogramming only occurs in 5-10% of cells exposed to reprogramming factors (Polo, Anderssen et al. 2012), it may not be feasible to clearly distinguish the differences even under longer doxycycline treatment. Thus, I will also take an alternative approach by immunohistochemistry of pluripotency markers such as alkaline phosphatase, SSEA-1, and Nanog. By examining the difference in relative abundance or presence of these markers will allow us to conclude whether the quiescent bulge cells indeed have higher genome plasticity.

Fourth, recent advances in assessing the kinetics of protein dynamics suggest that more thorough analyses of our FRAP data may be needed (Mueller, Mazza et al. 2010; Stasevich, Mueller et al. 2010; Mueller, Morisaki et al. 2012). First, discussions with Dr. Tom Misteli and Dr. Warren Zipfel, clearly suggested that simple diffusion of macromolecules including H2B-GFP takes place in the order of seconds. This indicates that the difference in the amount of recovery that takes place over the minutes (Figures 4.8G and 4.8J) is not likely due to the simple diffusion. However, the differences in the mobile fraction that we detect in both in *in vivo* and *in vitro* (Figure 4.8H and 4.8K) may be due to the differences in the total amount of H2B-GFP present in different cell types. In order to rule this out, I will perform a western blot of H2B-GFP extracted from keratinocytes at proliferative, quiescent, and differentiation conditions *in vitro*. Secondly, although many studies have shown similar results as ours (where the majority of the protein seems to be tightly associated to the chromatin with only a small diffusible component (Mueller, Karpova et al. 2012), different kinetic modeling can give varying outcomes of the dynamic state of the protein (Mueller, Karpova et al. 2012). In order to more quantitatively measure the kinetics, I will re-analyze the FRAP images (all cells and conditions represented in Figure 4.8I) using a different approach suggested by Dr. Warren Zipfel. I will take quantitative measurements of continuous strips of a few pixels covering the unbleached and bleached area. Just after the bleach, the edge of the unbleached and bleached area can be represented by a discontinuous function. However, as mobile fraction fills in the bleached area, the edge will bear sigmoid shape that can be represented by an error

function. Using this approach, I will be able to calculate more quantitatively the differences in how fast the H2B-GFP molecules migrate.

Lastly, although it is beyond the scope of this graduate work, to firmly establish the significance of epigenetic regulation of hair follicle stem cells quiescence and genome plasticity, genetic manipulation of histone modifying enzymes can be achieved. For instance, using a transgenic line expressing histone modifying enzyme(s) to over-express one or more specific methylases during quiescence will reveal how global epigenomic status influence the proliferative status as well as the genome plasticity. Conversely, similar approach can be used by ablation of specific demethylases during quiescence. Moreover, although our work only characterized a few of these, there are other histone modifications that specifically mark different genomic loci. Therefore, in the future, it will also be interesting to analyze the role of other histone marks such as acetylation or methylation (in other lysine residues) in hair follicle stem cells self-renewal and fate determination. In addition, considering that placement, maintenance, and removal of different histone modification marks involve other histone-associated proteins as well as chromatin remodeling complexes, it will also be interesting to determine how individual components play a role in a concerted action of epigenetic regulation in governing stem cells self-renewal and differentiation. Also, as multiple epithelial cancers such as squamous cell carcinoma are influenced by mis-regulation of hair follicle bulge cells self-renewal and differentiation, it will be of great importance to dissect how epigenetic factors contribute to tumorigenesis. Overall, the future directions documented here will significantly strengthen our model of how epigenomic state of hair follicle stem cells is regulated during quiescence and the

biological meaning of this mechanism in proper cell fate choice. The discoveries made here will potentially be applicable to other tissue stem cell systems and serve as a basis of understanding the molecular pathway of diseases such as cancer.

REFERENCES

Baxter, J., S. Sauer, et al. (2004). "Histone hypomethylation is an indicator of epigenetic plasticity in quiescent lymphocytes." EMBO J 23(22): 4462-4472.

Bernstein, B. E., T. S. Mikkelsen, et al. (2006). "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." Cell 125(2): 315-326.

Bracken, A. P., D. Kleine-Kohlbrecher, et al. (2007). "The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells." Genes Dev 21(5): 525-530.

Chen, J., H. Liu, et al. (2013). "H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs." Nat Genet 45(1): 34-42.

Cheng, T., N. Rodrigues, et al. (2000). "Hematopoietic stem cell quiescence maintained by p21cip1/waf1." Science 287(5459): 1804-1808.

Cotsarelis, G., T. T. Sun, et al. (1990). "Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis." Cell 61(7): 1329-1337.

de Nooij, J. C., M. A. Letendre, et al. (1996). "A cyclin-dependent kinase inhibitor, Dacapo, is necessary for timely exit from the cell cycle during Drosophila embryogenesis." Cell 87(7): 1237-1247.

Devgan, V., C. Mammucari, et al. (2005). "p21WAF1/Cip1 is a negative transcriptional regulator of Wnt4 expression downstream of Notch1 activation." Genes Dev 19(12): 1485-1495.

Ezhkova, E., W. H. Lien, et al. (2011). "EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair." Genes Dev 25(5): 485-498.

Ezhkova, E., H. A. Pasolli, et al. (2009). "Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells." Cell 136(6): 1122-1135.

Fasano, C. A., J. T. Dimos, et al. (2007). "shRNA knockdown of Bmi-1 reveals a critical role for p21-Rb pathway in NSC self-renewal during development." Cell Stem Cell 1(1): 87-99.

Ferrandiz, N., J. M. Caraballo, et al. (2012). "p21 as a Transcriptional Co-Repressor of S-Phase and Mitotic Control Genes." PLoS One 7(5): e37759.

Foudi, A., K. Hochedlinger, et al. (2009). "Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells." Nat Biotechnol 27(1): 84-90.

Frye, M., A. G. Fisher, et al. (2007). "Epidermal stem cells are defined by global histone modifications that are altered by Myc-induced differentiation." PLoS One 2(8): e763.

Fuchs, E. (2009). "The tortoise and the hair: slow-cycling cells in the stem cell race." Cell 137(5): 811-819.

Giadrossi, S., M. Dvorkina, et al. (2007). "Chromatin organization and differentiation in embryonic stem cell models." Curr Opin Genet Dev 17(2): 132-138.

Guertin, M. J. and J. T. Lis (2010). "Chromatin landscape dictates HSF binding to target DNA elements." PLoS Genet 6(9).

Hawkins, R. D., G. C. Hon, et al. (2010). "Distinct epigenomic landscapes of pluripotent and lineage-committed human cells." Cell Stem Cell 6(5): 479-491.

Hoi, C. S., S. E. Lee, et al. (2010). "Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin." Mol Cell Biol 30(10): 2518-2536.

Iwama, A., H. Oguro, et al. (2004). "Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1." Immunity 21(6): 843-851.

Jacobs, J. J., K. Kieboom, et al. (1999). "The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus." Nature 397(6715): 164-168.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science 293(5532): 1074-1080.

Kippin, T. E., D. J. Martens, et al. (2005). "p21 loss compromises the relative quiescence of forebrain stem cell proliferation leading to exhaustion of their proliferation capacity." Genes Dev 19(6): 756-767.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell 128(4): 693-705.

Lee, J. and T. Tumbar (2012). "Hairy tale of signaling in hair follicle development and cycling." Semin Cell Dev Biol.

Lien, W. H., X. Guo, et al. (2011). "Genome-wide maps of histone modifications unwind in vivo chromatin states of the hair follicle lineage." Cell Stem Cell 9(3): 219-232.

Luger, K., M. L. Dechassa, et al. (2012). "New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?" Nat Rev Mol Cell Biol 13(7): 436-447.

Margueron, R. and D. Reinberg (2011). "The Polycomb complex PRC2 and its mark in life." Nature 469(7330): 343-349.

Marques-Torrejon, M. A., E. Porlan, et al. (2012). "Cyclin-Dependent Kinase Inhibitor p21 Controls Adult Neural Stem Cell Expansion by Regulating Sox2 Gene Expression." Cell Stem Cell.

Matsumoto, A., S. Takeishi, et al. (2011). "p57 is required for quiescence and maintenance of adult hematopoietic stem cells." Cell Stem Cell 9(3): 262-271.

Meshorer, E., D. Yellajoshula, et al. (2006). "Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells." Dev Cell 10(1): 105-116.

Mikkelsen, T. S., M. Ku, et al. (2007). "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells." Nature 448(7153): 553-560.

Molofsky, A. V., S. G. Slutsky, et al. (2006). "Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing." Nature 443(7110): 448-452.

Mueller, F., T. S. Karpova, et al. (2012). "Monitoring dynamic binding of chromatin proteins in vivo by fluorescence recovery after photobleaching." Methods Mol Biol 833: 153-176.

Mueller, F., D. Mazza, et al. (2010). "FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know?" Curr Opin Cell Biol 22(3): 403-411.

Mueller, F., T. Morisaki, et al. (2012). "Minimizing the impact of photoswitching of fluorescent proteins on FRAP analysis." Biophys J 102(7): 1656-1665.

Onder, T. T., N. Kara, et al. (2012). "Chromatin-modifying enzymes as modulators of reprogramming." Nature 483(7391): 598-602.

Orford, K. W. and D. T. Scadden (2008). "Deconstructing stem cell self-renewal: genetic insights into cell-cycle regulation." Nat Rev Genet 9(2): 115-128.

Orkin, S. H. and K. Hochedlinger (2011). "Chromatin connections to pluripotency and cellular reprogramming." Cell 145(6): 835-850.

Osorio, K. M., S. E. Lee, et al. (2008). "Runx1 modulates developmental, but not injury-driven, hair follicle stem cell activation." Development 135(6): 1059-1068.

Pal, B., T. Bouras, et al. (2013). "Global changes in the mammary epigenome are induced by hormonal cues and coordinated by ezh2." Cell Rep 3(2): 411-426.

Park, I. K., D. Qian, et al. (2003). "Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells." Nature 423(6937): 302-305.

Pippa, R., L. Espinosa, et al. (2011). "p27(Kip1) represses transcription by direct interaction with p130/E2F4 at the promoters of target genes." Oncogene.

Polo, J. M., E. Anderssen, et al. (2012). "A molecular roadmap of reprogramming somatic cells into iPS cells." Cell 151(7): 1617-1632.

Richter, G. H., S. Plehm, et al. (2009). "EZH2 is a mediator of EWS/FLI1 driven tumor growth and metastasis blocking endothelial and neuro-ectodermal differentiation." Proc Natl Acad Sci U S A 106(13): 5324-5329.

Sangiorgi, E. and M. R. Capecchi (2008). "Bmi1 is expressed in vivo in intestinal stem cells." Nat Genet 40(7): 915-920.

Scheitz, C. J., T. S. Lee, et al. (2012). "Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer." EMBO J 31(21): 4124-4139.

Sen, G. L., D. E. Webster, et al. (2008). "Control of differentiation in a self-renewing mammalian tissue by the histone demethylase JMJD3." Genes Dev 22(14): 1865-1870.

Speck, N. A. and D. G. Gilliland (2002). "Core-binding factors in haematopoiesis and leukaemia." Nat Rev Cancer 2(7): 502-513.

Stadtfeld, M., N. Maherali, et al. (2008). "Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse." Cell Stem Cell 2(3): 230-240.

Stasevich, T. J., F. Mueller, et al. (2010). "Cross-validating FRAP and FCS to quantify the impact of photobleaching on in vivo binding estimates." Biophys J 99(9): 3093-3101.

Tesio, M. and A. Trumpp (2011). "Breaking the cell cycle of HSCs by p57 and friends." Cell Stem Cell 9(3): 187-192.

Topley, G. I., R. Okuyama, et al. (1999). "p21(WAF1/Cip1) functions as a suppressor of malignant skin tumor formation and a determinant of keratinocyte stem-cell potential." Proc Natl Acad Sci U S A 96(16): 9089-9094.

Tumbar, T. (2012). "Ontogeny and Homeostasis of Adult Epithelial Skin Stem Cells." Stem Cell Reviews and Reports 8(2): 561-576.

Tumbar, T., G. Guasch, et al. (2004). "Defining the epithelial stem cell niche in skin." Science 303(5656): 359-363.

van Os, R., L. M. Kamminga, et al. (2007). "A Limited role for p21Cip1/Waf1 in maintaining normal hematopoietic stem cell functioning." Stem Cells 25(4): 836-843.

Velichutina, I., R. Shaknovich, et al. (2010). "EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis." Blood 116(24): 5247-5255.

Viale, A., F. De Franco, et al. (2009). "Cell-cycle restriction limits DNA damage and maintains self-renewal of leukaemia stem cells." Nature 457(7225): 51-56.

Waghmare, S. K., R. Bansal, et al. (2008). "Quantitative proliferation dynamics and random chromosome segregation of hair follicle stem cells." EMBO J 27(9): 1309-1320.

Wilson, N. K., S. D. Foster, et al. (2010). "Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators." Cell Stem Cell 7(4): 532-544.

Ye, L., Z. Fan, et al. (2012). "Histone demethylases KDM4B and KDM6B promotes osteogenic differentiation of human MSCs." Cell Stem Cell 11(1): 50-61.

Yu, M., T. Mazor, et al. (2012). "Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors." Mol Cell 45(3): 330-343.

Zhang, Y. V., J. Cheong, et al. (2009). "Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells." Cell Stem Cell 5(3): 267-278.

Zhang, Y. V., B. S. White, et al. (2010). "Stem cell dynamics in mouse hair follicles: a story from cell division counting and single cell lineage tracing." Cell Cycle 9(8): 1504-1510.

Zou, P., H. Yoshihara, et al. (2011). "p57(Kip2) and p27(Kip1) cooperate to maintain hematopoietic stem cell quiescence through interactions with Hsc70." Cell Stem Cell 9(3): 247-261.